Dealing with Rising Health Care Costs: The Case of Pharmaceuticals

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

im Jahr 2006
an der Volkswirtschaftlichen Fakultät der
Ludwig-Maximilians-Universität München
vorgelegt von

Ingrid Susanne Königbauer

Referent: Prof. Ray Rees
Korreferent: Prof. Dr. Andreas Haufler
Datum der mündlichen Prüfung: 12. Juli 2006
Promotionsabschlussberatung: 26. Juli 2006

Acknowledgements

This doctoral thesis was written under the supervision of Professor Ray Rees, University of Munich, from 2002 to 2006. I would like to thank Professor Ray Rees for his openness to my ideas which allowed me to follow my interests without any restrictions. I am glad for his help with earlier drafts and his numerous comments and suggestions. It was a pleasure to work at his chair where he managed to keep the non-scientific workload as low as possible.

The good atmosphere at the chair is also due to my great colleagues. I would like to thank Tobias Böhm, Florian Englmaier, Elisabeth Meyer, Astrid Selder, and Hans Zenger with whom it was a real pleasure to work. A special thanks goes to Astrid who influenced me a great deal in my choice to pursue health economics – a choice which I do not regret in retrospect – and to Tobias whose constructive critique had a great impact on my work.

I would also like to thank Irmgard von der Herberg who handled most of the administrative work and alleviated life for me at the chair immensely. Many student helpers assisted me in scientific and administrative matters. I would like to mention especially Florian Bitsch, Kinga Funk, Nico Groz, and Max von Liel.

Parts of this doctoral thesis were written during my research visit at the University of Bergen, Norway, with the financial support of the Deutscher Akademischer Auslandsdienst (DAAD) and the Ruhrgas grant. I am grateful to both organisations. I would also like to thank Jan Erik Askildsen as the head of the HEB for enabling me to come and for taking so good care of me. I am glad for having had the opportunity to work with Kurt Brekke who not also were willing to discuss my papers and ideas with me, but also tried and managed to make my stay very pleasant. Additionally, I want to thank Odd Rune Straume - with Odd Rune and Kurt, I was writing the paper Reference Pricing of Pharmaceuticals which is the basis for chapter 3 of this doctoral thesis.

Furthermore I thank Professor Dr. Andreas Haufler and Professor Dr. Monika Schnitzer who kindly agreed on joining the board of referees for this thesis.

Finally, I thank my parents who supported me during all my education and Philipp who patiently endured all major and minor crises during my PhD-studies.

March 2006 Ingrid Königbauer

Contents

Chapter 1

Introduction

Already in 1958, Mushkin portraits nicely the dilemma that most industrialised countries nowadays face and which gets increasingly more problematic: On the one hand, there are the sick and ill who need expensive treatment, on the other hand, resources are scarce and need to be allocated efficiently. This is a classic trade-off that economists are trained for to solve and that makes the discipline health economics more and more important in the future.
The beginning of health economics as an individual discipline within economics dates in general with the classic article of Arrow (1963) in which he analyses the medical care industry and its (in-)efficiency. Medical care differs from the ordinary commodity in several aspects which are defined by Hudson (2000) as follows: Demand for health care is a derived demand for health, i.e. an input to produce or restore the commodity 'health', externalities, informational asymmetries between providers and patients, and uncertainty with respect to both the need for and the effectiveness of health care. Though each one of the aspects can be encountered in other markets as well, it is the simultaneous incidence which makes the medical market so challenging for thorough health policy, whose main purpose should be the provision of circumstances that allow for an efficient production and distribution of medical resources.

Good health is in general seen as one of the most valuable goods that a human being can achieve. The wealthier people get, the more they are willing to spend in order to restore the state of health in case of a more or less restricting illness. But also the technologies

and treatments become increasingly advanced and consequently more and more expensive. Together with an ageing population, this leads to dramatic increases in the health expenses worldwide. Within the OECD countries, the average expenses for health doubled from about 5 percent of the GDP in 1960 to about 10 percent of the GDP in 2000 (Breyer et al. (2003)). This increase in health expenses is dramatic and likely to continue, if no further measures are taken.

Pharmaceuticals are important aspects of the provision of health care and are the second highest expense block for the compulsory health insurance in Germany behind the expenses for stationary treatment and yet higher than the ambulant treatment costs (Sachverständigenrat (2005)). Hence, pharmaceuticals constitute a respectable proportion of total expenditure in health care. Together with the seemingly exorbitant high profits that the pharmaceutical industry enjoys, this leads to the understandable claim that it contributes heavily to the 'crisis in the health sector'. This is mainly due to the special features of the pharmaceutical market that result in a low cost-consciousness of the consumers and high market power of the suppliers:
On the demand side, uninformed consumers delegate the treatment decision to experts (physicians) whose preferences might not be perfectly aligned with the consumers' preferences. Additionally, prescription drugs are often reimbursed at least partly by insurance which decreases the price-elasticity of demand.
On the supply side, innovative drugs are temporarily protected by a patent after the introduction to the market. Patents are necessary for the pharmaceutical firms in order to recoup the huge R&D-costs that are incurred in order to produce welfare-enhancing innovative drugs. During this patent protection, the supply side is characterised by monopolistic competition which leads to higher prices. The only competition, that occurs during patent protection, is the competition with me-too drugs that are therapeutically equivalent and therefore close substitutes. After patent expiry, there is potential price competition with and between generic substitutes. A generic drug must prove to be bioequivalent to the original drug and is therefore objectively a perfect substitute to the original version.[1] However, due to reputational, habitual, or security reasons, as well as artificial product differentiation due to vigorous marketing efforts, the substitutability between original brand-name

[1]Hudson (2000): Two drugs are said to be bioequivalent, if they are chemically identical and have an approximately equal bioavailability. The concept of bioavailability relates to how much of the drug's active ingredients get into the bloodstream, and to the site and rate of therapeutic action.

drugs and its generic substitutes is less than perfect and therefore competition between them limited.

The extent to which generic versions firstly enter the market and secondly are substituted for brand-name drugs varies widely from nation to nation[2], and it remains an important question for every country, how the use of generics can be spurred further. Many initiatives came from the demand side by increasing the cost-consciousness of both the physicians and the patients – physicians becoming subject to budget limits for pharmaceuticals, and patients facing differential copayments. Additionally, the rise of cost-conscious health care organisations like the Managed Care Organisations in the United States are supposed to increase the use of generic substitutes.
But the use of generics can also be affected by implementing appropriate conditions on the supply side. Chapter 2 deals with this problem by emphasising one strategy, namely brand-name advertising, and its effect on the generic market entry decision.

The weak price competition both during patent protection, but especially after patent expiry, is the reason why most countries resort to price regulation in order to curb pharmaceutical prices and thus health costs. Ballance et al. (1992) surveyed 56 nations with respect to their regulations of the pharmaceutical industry and found out that 30 nations can be classified as having substantial price regulations, 20 countries have limited price regulations, and only 6 countries do not regulate their prices. The variety of measures to regulate prices is vast and they can be classified roughly as follows:
Supply-side interventions are price ceilings, negotiated price-volume agreements, and profit caps. With price ceilings like in France, Italy, and Spain, every individual drug price must be approved, if it is supposed to be reimbursed by the national health insurance. Australia, Austria, and France apply negotiated price-volume agreements, which are based on sales forecasts and where prices need to be reduced as compensation in case of a volume overage. In the United Kingdom, the return on capital of those pharmaceuticals reimbursed by the NHS cannot surpass the ex-ante determined profit cap.
Demand-side interventions are pharmaceutical budgets for physicians, positive and negative lists, and patients' out-of-pocket payments. With budgets, physicians face specific limits for pharmaceuticals that they are not allowed to surpass without risking personal liability.

[2]Germany's market share of generics with 30 percent of the national health insurance's overall pharmaceutical expenses is supposedly the highest worldwide (Sachverständigenrat (2005)).

This has been temporarily implemented in Germany. With positive and negative lists, a drug must either be positively approved for reimbursement, or drugs which are not approved are put on a negative list. Additionally, parts of the pharmaceutical costs are passed on to patients in order to make them more cost-conscious. Examples for patients' out-of-pocket payments are copayments or reference pricing.
Chapter 3 focuses on reference pricing as an indirect way to regulate pharmaceutical prices. In Germany, the Netherlands, Denmark, and New Zealand among others, pharmaceuticals are grouped together into clusters. A reference price is allocated to each group which determines the maximum price reimbursed by the national health insurance. The question is emphasised, how different definitions of similar drugs affect the resulting drug prices and welfare.

Although price regulation seems to succeed in the short-term goal to reduce pharmaceutical costs, it might entail negative long-term impacts on pharmaceutical innovation. It is well known that the pharmaceutical industry is a very R&D-intensive industry. Danzon (1997b) calculates that R&D accounts for about 30 percent of total costs. These sunk costs must be recovered in order to maintain sufficient incentives to innovate.[3] Numerous studies have found a significant positive link between profitability, affected by price regulations, and innovation activity, e.g. Gambardella (1995) and Grabowski and Vernon (2000), and accordingly the drug development, e.g. Acemoglu and Linn (2003). Although many innovations might only constitute minor developments which express themselves in reduced side effects or higher patients' compliance, the majority of the literature comes to the conclusion that, in general, new medications reduce the morbidity and the length and number of hospital stays, and that they increase longevity (see e.g. Lichtenberg (2001)). These better health outcomes enhance indirectly a country's economic standing by improving the workforce.
However, the benefits of new drug developments can only be enjoyed, if the incentives to innovate are maintained. In this context, it is often argued that this can become a problem in the international context, if prices are regulated and if parallel imports are allowed (Danzon (1997b)), as it is explicitly the case among the countries within the European Union. Parallel imports are legitimately produced drugs which are imported into another country legally, but without the authorisation of the patent holder, with the aim to exploit arbitrage possibilities. Chapter 4 deals with this aspect and shows that the threat of abolished R&D-

[3]See Hassett (2004) for a survey on the effects of price controls on pharmaceutical markets.

incentives due to marginal cost pricing, as it is portrayed in the literature, is exaggerated. This conclusion is based on the fact that pharmaceutical firms can anticipate potential parallel imports and credibly threat not to supply individual countries, if their price regulation is too strict.

In the following, a brief summary of the chapters 2 to 4, which are based on individual essays, is provided. The motivation and a more extensive discussion of the literature will be provided in each chapter.

Chapter 2 analyses the relationship between Advertising and Generic Market Entry, i.e. the effect of brand-name advertising on off-patent price competition. Advertising on the pharmaceutical market is extensive and the expenses for advertising easily reach those for R&D (Hurwitz and Caves (1988), Sachverständigenrat (2005)). Detailing, i.e. promotional activities targeted at physicians is thereby the highest share of the expenses. It is in general allowed, because it is believed that the positive welfare effect due to the provided information dominates the negative welfare effect due to persuasion and the resulting brand-loyalty. This chapter, however, emphasises the (negative) persuasive aspect of advertising and shows that, even then, advertising per se is not a barrier to market entry.
By investing in advertising, the brand-name firm is able to build up a stock of goodwill, i.e. physicians who either believe the original drug's quality to be higher or who feel obliged to the pharmaceutical firm. After patent expiry, a generic version of the drug enters the market, if the anticipated generic profit can cover the fixed market entry costs. Due to the perceived product differentiation induced by detailing, competition is softened and some positive profits are possible for the entrant. This advantage of advertising due to product differentiation can dominate the negative effect due to brand-loyalty, if an upper limit of advertising is not surpassed.
In this context, it is interesting whether advertising can be beneficial from a social point of view, even though it might not offer any informational content. Advertising can increase welfare, because it induces generic market entry and leads to lower off-patent prices. The downside of advertising is that, on the one hand, it increases on-patent prices and it incorporates wasteful advertising costs, and on the other hand, generic market entry can be deterred by over-investing in advertising. Ex ante, it is not obvious which of the effects dominates and it can be shown that even purely persuasive advertising can be socially beneficial for a sufficiently effective advertising technology, a short period of patent protection,

and small generic market entry costs. A benevolent regulator should try to affect welfare positively by manipulating the conditions under which the pharmaceutical firm chooses the profit-maximising advertising level:
A sufficiently high degree of product differentiation between original drugs and generic drugs is necessary in order to induce market entry. It can be influenced by providing objective information about the quality of drugs and by regulating the extent to which physicians can be bribed with presents and sponsored conference trips.
When choosing the optimal length of patent protection, not only the R&D-incentives should be considered, but also its effect on the amount of advertising. This is particularly important, because the expenses for advertising can easily reach those for R&D.
The direct setup-costs of generic firms cannot be affected by the health authorities. But by reducing the bureaucracy and formal requirements associated with generic market entry, the likelihood of the introduction of generic substitutes can be increased.
As an extension of the model, it is analysed in the same framework, how price regulation affects generic market entry in the presence of (persuasive) advertising. The model confirms the empirical finding that generic market shares are lower in countries where price regulation is stricter (Danzon and Chao (2000)), because the potential generic entrant expects profits that are the lower, the stricter prices are regulated. Hence, there is a trade-off with respect to welfare: On the one hand, strict price regulation reduces prices and increases simultaneously (patients') welfare. On the other hand, price regulation makes it more likely that generic market entry is deterred and that there is no lower-priced drug version available.

This chapter is based on the following two working papers that I worked on during my research visit at the University of Bergen: "Advertising and Generic Market Entry", HEB Working Paper 03/05, and a revised version of HEB Working Paper 04/05, "Advertising, Generic Competition, and Price Regulation in Pharmaceutical Markets". In 2004, earlier versions were presented at seminars of the University of Munich and the University of Bergen, at the CESifo Summer Workshop "Health Economics" in 2005, and at the 32nd Conference of the European Association for Research in Industrial Economics 2005. These papers benefited greatly from discussions with Kurt Brekke, Tobias Böhm, Normann Lorenz, Ray Rees, Astrid Selder, and Odd Rune Straume. A special thanks goes to Helge Berglann who helped me with the numerical simulations.

In chapter 3, one specific type of price regulation, Reference Pricing, is analysed in

greater detail. Reference pricing is quite novel, but has rapidly become a widely used price control mechanism in the pharmaceutical market since its first introduction in Germany in 1989. Under a reference price system, drugs are classified into clusters based on similar therapeutic effects and a reference price is attributed to each cluster. The reference price is the maximum reimbursement for all drugs in the group. The patient must pay the surcharge, if the drug's price exceeds the reference price level. Several countries have already established a reference price system, but they differ in the exact design. Whereas the international differences in the exact level of the reference price and the copayments to be borne by the patients are minor such that they have only quantitative, not qualitative consequences, the difference in the definition of 'similar' drugs is severe: If many drugs are combined under the same reference price, i.e. also on-patent, horizontally differentiated me-too drugs, as this is done under a therapeutical reference price system, then the drug prices as well as the difference between the drug prices can be reduced as compared to the more lenient definition under the generic reference price system, where only off-patent drugs and their generic substitutes are included. A strict reference price system can reduce the expenses for pharmaceutical drugs *and* the patients' loss due to mismatch costs that arise, because patients differ in their susceptibility towards various therapeutically equivalent drug versions. However, these benefits come at the expense that new treatments might refrain from entering the market with a strict reference price regime. The question, that is addressed in this chapter, is therefore, how the different types of reference price systems affect the availability of different drug versions and patients' welfare, and why different types of reference price systems are implemented in various countries.

From the model, no clear-cut conclusions can be drawn about the optimal choice of the reimbursement system. However, the following classification can be derived: Therapeutical reference pricing minimises both the patients' mismatch costs and the drug prices. It is therefore clearly the socially favourable reimbursement system, if the risk of no market entry for new drugs is low. However, if this is not the case, then either no reference pricing or generic reference pricing might be necessary to stimulate market entry. Free pricing minimises the mismatch costs, but maximises the drug prices. It should be found in countries where drug prices do not play an important role for social welfare, which might be due to a dominant pharmaceutical industry. Generic reference pricing maximises the patients' mismatch costs, but at lower prices than under free pricing. It might be favoured in countries, where the pharmaceutical industry is insignificant or non-existent.

This chapter is based on a joint paper with Kurt Brekke and Odd Rune Straume, whom I both met during my research visit at the University of Bergen. Earlier versions were presented in seminars at the NHH in Bergen and at the University of Munich in 2005.

In chapter 4, the relationship between Parallel Imports and Price Regulation is analysed. The argument of Danzon (1997b) is investigated that governments have an incentive to free-ride on the huge sunk pharmaceutical innovation costs by setting drug prices equal to marginal costs. This is individually rational, because it minimises health costs, but it reduces or eliminates the pharmaceutical firms' incentives to invest in R&D and leads to a suboptimal supply of new products. This result hinges, however, on the fact that a closed economy is considered where parallel imports are not allowed.
The situation in the European Union is considered, where parallel imports are explicitly allowed, and it is shown that there is no marginal cost pricing in this context. Countries have an incentive to regulate their pharmaceutical prices as low as possible, but they also need to induce the pharmaceutical firms to export the products to their countries. They need to consider the possible negative spill-over effect of too low prices: Pharmaceutical firms anticipate that low prices will be reimported into high-priced countries. This can result in lower overall profits than without the additional low-priced country. Therefore, the regulated price ceiling has to be sufficiently high in order to guarantee market coverage with that drug. With the number of foreign countries within the parallel-importing area sufficiently high, all foreign countries will set a price ceiling only slightly above marginal costs, thus taking advantage of the quantity effect – all foreign countries together need to fulfil the participation constraint. The mark-up on marginal costs will be determined by the home country, which sets the highest possible ceiling, i.e. the monopoly price. The pharmaceutical firm will make monopoly profits based on the demand at home.
Several interesting policy implications can be drawn from the model: The model strongly supports the European policy to explicitly allow parallel imports and it shows that it is less severe to delegate price regulation to the individual national governments than it is often assumed. Even the expansion of the European Union by several Eastern European countries is less threatful as commonly thought, because the firm's profits should not be affected.
Parallel imports should not be prevented, but rather actively fostered with as many countries as possible. Including, e.g., developing countries into the parallel-importing area, which will cause a large quantity effect, will lead to lower acceptable price ceilings. Eventually, also poor countries might be able to afford necessary drugs.

The model predicts that R&D is located in high-income and low-price-elasticity countries, because then the highest possible profit can be guaranteed by the home government. Although feared to introduce further losses to the pharmaceutical industry, the model shows that the regulatory requirements, which were harmonised by the European Medicines Evaluation Agency in 1996, do not affect profits, but rather increase consumer welfare.

In 2004, earlier versions of this paper were presented in seminars at the University of Munich, at the Spring Meeting of Young Economists, at the 19th Meeting of the European Economic Association, and the 2nd Symposium on "Gesundheitsökonomische Grundlagen für die Gesundheitspolitik", organised by the DIW Berlin and the Hans Böckler Stiftung. The study benefited greatly from discussions with Tobias Böhm, Florian Englmaier, Ray Rees, and Astrid Selder. A non-technical version in German, that focuses more on policy implications, is published under the title "Die Auswirkung von Parallelimporten auf die optimale Regulierung von Arzneimittelpreisen" in the Vierteljahrshefte zur Wirtschaftsforschung 2004(4), 592-604.

Chapter 5 concludes. All technical proofs are relegated to the Appendix in chapter 6 and the References can be found in chapter 7.

Chapter 2

Advertising and Generic Market Entry

2.1 Introduction

Since patients are uninformed and lack the information about which treatment is most effective, they depend on physicians who diagnose and suggest some treatment. Thus, physicians directly affect the extent of competition between different providers of a treatment and can be taken as the main determinant of whether a brand-name or a generic drug version is prescribed, as the empirical evidence in Hellerstein (1998) suggests. Hence, it is not surprising that in the pharmaceutical market, where price competition would be fierce, if the products were not differentiated somehow, the physician is the target of huge advertising expenditures. Advertising expenditures with about 20-30 percent of sales are often even larger than those for R&D (Hurwitz and Caves (1988)). Jacobzone (2000) shows that within the OECD countries, the research-oriented pharmaceutical firms spent 24 percent of sales on marketing in 1989. And Scherer (2000) reports that in the United States, the ethical pharmaceutical industry spent 18 percent of sales on marketing in 1997.

Despite being rather old, these studies fit well with recent data concerning the pharmaceutical industry's expenditures on marketing as compared to R&D expenses. A recent study by the German regulatory authority for health, Sachverständigenrat (2005), confirms that in 2004, the ten major innovative pharmaceutical firms spent between 29 % and 38 % of sales

on marketing and administration, as compared to between 13 % and 19 % on R&D.

Advertising in the ethical pharmaceutical market, i.e. the market for prescription drugs, is in general allowed, if targeted towards experts (the physicians) on the grounds that it provides necessary information which might dominate the downsides of advertising.[1] This paper, however, emphasises the persuasive aspect of advertising and thus the 'negative' aspect of advertising and shows that, as in the literature on informational advertising[2], advertising per se is no barrier to market entry. The advantage of persuasive advertising, namely product differentiation which induces generic market entry and thus lower post-patent prices, can dominate the negative effect due to brand-loyalty, if an upper limit of advertising is not surpassed. It is analysed under which circumstances generic market entry is most likely and how these conditions can be positively influenced by the health authorities.

There is a rather extensive empirical literature that tries to capture the effect of advertising on generic market entry.[3] Hurwitz and Caves (1988) find that current and past investments in goodwill (advertising) preserve the incumbent's market share, whereas generic price discounts reduce it. Rizzo (1999) finds that brand-name advertising decreases the price-elasticity of demand in the pharmaceutical industry, because it increases brand-loyalty. Based on these effects, both papers conclude that brand-name advertising inhibits generic market entry. The present model incorporates both the goodwill effect of advertising and the decreased price-sensitivity. Similarly, it reaches the conclusion that generic market entry can be deterred by over-investing in advertising. However, since advertising induces vertical product differentiation, some advertising is a necessary prerequisite for generic market entry. As long as the incumbent invests sufficiently little in advertising, this positive differentiation effect dominates the negative brand-loyalty effect.
Scott Morton (2000) examines the role of on-patent brand-name advertising on the generic market entry decision after patent expiry. She thereby assumes advertising to be an endogenous variable which might be used to deter generic market entry. She finds that advertising

[1] In the United States, pharmaceutical advertising targeted at the patients is also allowed and a much debated issue. This analysis focuses, however, on detailing, i.e. on advertising targeted at physicians, and does not consider direct-to-consumer advertising.

[2] See Schmalensee (1983), Ishigaki (2000), and Fudenberg and Tirole (1984).

[3] See also the following rather old studies: Vernon (1971) finds no statistically significant effect of advertising on market entry. Telser et al. (1975) and Leffler (1980) both find a positive relationship between advertising and generic market entry.

is not a barrier to market entry.

Theoretically, advertising in the pharmaceutical market has mostly been modelled with respect to competition between therapeutically equivalent brand-name drugs, i.e. horizontal product differentiation models have been applied.[4] There are some notable exceptions. Frank and Salkever (1992) analyse the brand-name pricing behaviour after generic market entry in the presence of advertising. In a Stackelberg setup, they derive the conditions under which the model can explain the empirical finding of minimal brand-name price decreases or even increases after generic market entry and the sharp decline in brand-name advertising. This appears to be especially the case, if generic market entry leads to a significant decline in the price-elasticity of the brand-name demand, and if advertising leads to a substantial differentiation between brand-name and generic drug versions.
Cabrales (2003) studies oligopolistic competition in off-patent pharmaceutical markets, where advertising is used to create perceived differences in the quality of the brand-name and generic drug versions. However, he does not focus on the generic market entry decision, but rather analyses the effect of price regulation on generic market shares and overall quality provision.

Equivalent to Cabrales (2003), the present model uses vertical product differentiation between brand-name and generic drugs. But it focuses on another aspect: It stresses that advertising is a way to induce the necessary differentiation between brand-name and generic firms that makes generic market entry possible and simultaneously reduces off-patent prices. Generic market entry is deterred without advertising, because this situation would lead to detrimental Bertrand competition that does not allow the generic firm to recover the market entry costs. However, there is an upper limit of advertising above which the negative effect due to brand-loyalty dominates the positive differentiation effect.
The situation is modelled in a 2-periods-game. In the first period, the incumbent is protected by a patent and invests in advertising targeted at the prescribing physicians (detailing).[5] Advertising creates two markets and the detailed physicians perceive the brand-name quality

[4]Brekke and Kuhn (2006), e.g., analyse the interaction between detailing and direct-to-consumer advertising.

[5]Detailing means that the pharmaceutical firm's representative visits physicians in order to advertise a specific drug. Detailing with about 70 percent (Hurwitz and Caves (1988)) has the biggest share of advertising targeted at the physician. Additionally, advertising targeted at physicians also encompasses printed advertising, sponsored conferences, and so on.

to be higher than the not-detailed physicians. With this additional vertical differentiation, the incumbent can reduce competition after patent expiry. In the second period, there is potential generic market entry. Since market entry is costly, the expected generic profit must be sufficiently high. With respect to advertising, market entry can be induced by differentiation, but it can just as well be deterred by brand-loyalty.

In an extension of the model, I analyse in the same framework, how price regulation affects generic market entry in the presence of advertising. The empirical finding is confirmed that generic market shares are lower in countries where price regulation is stricter (Danzon and Chao (2000)), because the potential generic entrant expects profits that are the lower, the stricter prices are regulated. Hence, there is a trade-off with respect to welfare. On the one side, strict price regulation reduces pharmaceutical prices directly. On the other side, however, price regulation can deter generic market entry and thus inhibit an even lower-priced drug version.

The chapter is structured as follows: In section 2.2, the model is presented. The case without price regulation is analysed in section 2.3. In section 2.4, the effect of price regulation on generic market entry is discussed. Finally, section 2.5 concludes.

2.2 The Model

Assume that the market for prescription drugs consists of a continuum of patients distributed uniformly on the segment $[0, \bar{t}]$, and the patients suffer from the same illness which can be treated with a brand-name drug B and potentially with a generic drug G. There are two periods: During patent protection, only a brand-name drug is available, whereas after patent expiry, a generic substitute might be offered. The patients' position t can be interpreted as the extent to which they are ill and corresponds to their valuation for treatment $v(t) = t$, with $v'(t) = 1 > 0$, i.e. the more severely ill the patients are, the higher is their willingness to pay for a drug. The utility of a patient located at t from getting the drug i is thus

$$U_t^i = t - p^i \tag{2.1}$$

where p^i is the price of drug $i = B, G$. Note that the patients bear the whole drug expenses themselves by assumption, and not only a copayment as it is mostly the case in practice. The model and analysis does not change qualitatively, if such a copayment was introduced.

The patients are not informed about the available medication and they therefore need to go to a physician to get a suitable prescription. There is a mass of $N = 1$ ex ante identical physicians among whom they can choose and who prescribe either the branded or the generic version of the drug. Since the generic drug had to prove bioequivalence, it can be assumed that both the brand-name and the generic version have the same quality normalised to 1. But physicians can be the targets of brand-name advertising. Brand-name advertising in this model is purely persuasive and does not contain any valuable information. The only aim of advertising is to distort the physicians' prescription behaviour in favour of the brand-name drug. The detailed physicians' perceived quality of the brand-name drug increases from t to θt, with $\theta > 1$, whereas a not-detailed physician attaches the same valuation t to both drug versions.[6] Alternatively, physicians can be seen as being corrupted or obliged to the pharmaceutical company to the extent of θ. Hence, advertising leads to vertical market segmentation, but it does not affect the extent to which the quality perception increases.

The physicians base their prescription choice on the perceived patient's utility and they are assumed to observe the patient's individual valuation t perfectly. If they are not detailed, they maximise the patient's utility as stated in (2.1), whereas they maximise the distorted utility, if they are detailed:

$$U_t^i = \begin{cases} \theta t - p^B & \text{if} \quad i = B \\ t - p^G & \text{if} \quad i = G \end{cases} \tag{2.2}$$

This assumption is in line with the empirical finding of Lundin (2000) that the physicians care both about the patients' valuation for a specific drug and their expenditures.

The brand-name firm decides on the fraction k of physicians that is supposed to be detailed and whose valuation increases from t to θt at the cost

$$A(k) = \frac{1}{\gamma + 1} k^{\gamma+1} \tag{2.3}$$

with $\gamma > 1$, $A'(k) = k^\gamma > 0$, and $A''(k) = \gamma k^{\gamma-1} > 0$. Thus, the cost of detailing is assumed to be increasing in the fraction of physicians to be reached and convex. This is a

[6]This is in line with articles in the medical literature: Fridman et al. (1987) found out, for example, that only half of 245 surveyed physicians believe that generic drugs are as effective as the original. Avon et al. (1982) conclude that the physicians' belief about certain drugs are more in line with the advertisement claims than actual measures of the pharmaceutical's performance. Chren and Landefeld (1994) found out that the likelihood for a physician to request a specific drug to be added to a hospital's formulary depends positively on his interaction with the drug company.

standard cost function in this strand of literature (see e.g. Cabrales (2003)). Note that due to $k \in [0,1]$, the advertising costs are smaller, the larger is γ. It is assumed that advertising is perfectly remembered by the physicians, i.e. there is no depreciation, and that only the brand-name firm advertises. This is in line with empirical findings that generic advertising in the prescription drug market is practically non-existent (Scherer (2000)).

After detailing, the market of physicians is divided into a detailed and a not-detailed market, where no price discrimination between detailed and not-detailed physicians is possible.[7] Since both types of physicians base their prescription behaviour on the patients' utility as stated in (2.1) and (2.2), it differs according to whether they are detailed or not due to differences in the perceived quality. In the detailed market, if there is no generic alternative available, the physicians prescribe the brand-name drug to all patients with a sufficiently high valuation for treatment:

$$\theta t - p^B \geq 0 \quad \Leftrightarrow \quad t \geq \frac{1}{\theta} p^B \tag{2.4}$$

If there is a generic alternative available, a vertical product differentiation model as in Gabszewicz and Thisse (1979) and Shaked and Sutton (1982) must be applied. The physicians prescribe the brand-name drug to all patients whose valuation for treatment is larger than $\hat{t}$:

$$\theta t - p^B \geq t - p^G \quad \Leftrightarrow \quad t \geq \hat{t} := \frac{p^B - p^G}{\theta - 1} \tag{2.5}$$

It will be shown later in the analysis that $p^B > p^G$ in equilibrium. All other patients with a sufficiently high valuation for treatment receive the generic drug, whereas the patients with little valuation for treatment receive no drug:

$$t - p^G \geq 0 \quad \Leftrightarrow \quad t \geq p^G \tag{2.6}$$

The not-detailed physicians assume the correct brand-name quality. If there is no generic alternative available, they prescribe the brand-name drug to all patients with a sufficiently

[7] Theoretically, the incumbent might find it optimal to introduce a branded generic drug, i.e. the same drug under a different label and at a lower price, in order to price-discriminate. Scherer (2000) states, however, that most brand-name producers do not find this strategy profitable, because too many patients would be lost to the lower-priced version. This will be assumed in this model. The results would not change qualitatively, however, if this simplification was not made, because the incumbent would only introduce a generic substitute of his own, if another generic firm threats with market entry.

high valuation for treatment:

$$t - p^B \geq 0 \quad \Leftrightarrow \quad t \geq p^B \tag{2.7}$$

With an available generic alternative, there will be fierce Bertrand competition, because the physicians do not believe that the two drug versions differ in quality. The not-detailed physicians exclusively prescribe the least expensive (generic) drug to all patients with a sufficiently high valuation for treatment:

$$t - p^G > t - p^B \quad \Leftrightarrow \quad p^G < p^B \tag{2.8}$$

$$t - p^G \geq 0 \quad \Leftrightarrow \quad t \geq p^G \tag{2.9}$$

The patients are assumed to be ignorant as to whether a physician has been detailed or not. Since they cannot observe every single contact between a physician and a pharmaceutical firm's representatives, they decide randomly which physician to visit. Therefore, every physician has the same representative sample of patients uniformly distributed on $[0, \bar{t}]$.

Based on this prescription behaviour, the demand functions can be derived taking into account the vertical market segmentation generated by advertising k. The demand functions differ depending on the period and whether there is generic market entry. In the first period, there is only the brand-name firm in the market. Its demand is therefore

$$D_1^B(p_1^B, k) = k\left(\bar{t} - \frac{1}{\theta}p_1^B\right) + (1-k)\left(\bar{t} - p_1^B\right) \tag{2.10}$$

Both detailed and not-detailed physicians prescribe the brand-name drug, but demand is larger in the detailed market due to the higher perceived quality.
In the second period, it must be differentiated between the situation without generic market entry and the corresponding brand-name demand as in (2.10), and the situation with generic market entry:

$$D_2^B(p_2^B, p_2^G, k) = k\left(\bar{t} - \hat{t}\right) \tag{2.11}$$

$$D_2^G(p_2^B, p_2^G, k) = k\left(\hat{t} - p_2^G\right) + (1-k)\left(\bar{t} - p_2^G\right) \tag{2.12}$$

Since the brand-name price is higher than the generic price, the brand-name drug is only prescribed to the high-valuation patients in the detailed market, whereas the generic drug is prescribed to the treated low-valuation patients in the detailed market and to all treated patients in the not-detailed market. Any change in the price level or the perceived quality

affects the demand, because the indifferent patient $\hat{t}$ changes. Demand for the brand-name firm decreases with the own price level and increases with the generic price level. The larger the perceived quality difference is, the larger is the incumbent's demand. Additionally, it weakens the own negative price effect on demand, whereas it reinforces the competitor's positive price effect on own demand. Demand for the generic firm reacts just the opposite way.

It is assumed that both firms face identical and constant marginal costs normalised to zero, but the generic entrant faces additionally some positive fixed market entry costs F which are sunk after market entry has occurred. These fixed market entry costs can be interpreted as costs which are due to the necessary market research that a generic firm has to undertake before the drug launch. In order to define the optimal pricing strategy, the generic firm has to analyse, e.g., the market structure, the competitors, and the potential demand. Furthermore, each national regulatory agency might need additional information for its national (price) regulation purposes.

In order to compare the length of patent protection with the period length without patent protection, a factor δ is introduced as a weight for period 1 respectively $(1-\delta)$ for period 2. The larger δ, the longer patent protection is compared to the period without patent protection.

The brand-name firm's overall profit without generic market entry can then be written as

$$\pi^B(p^B,k) = \left[k\left(\bar{t}-\frac{1}{\theta}p^B\right)+(1-k)\left(\bar{t}-p^B\right)\right]p^B - \frac{1}{1+\gamma}k^{1+\gamma} \tag{2.13}$$

With a generic substitute, both firms' overall profits are

$$\pi^B(p_1^B,p_2^B,p_2^G,k) = \delta\left[k\left(\bar{t}-\frac{1}{\theta}p_1^B\right)+(1-k)\left(\bar{t}-p_1^B\right)\right]p_1^B + (1-\delta)\left[k\left(\bar{t}-\hat{t}\right)\right]p_2^B - \frac{1}{1+\gamma}k^{1+\gamma} \tag{2.14}$$

$$\pi^G(p_1^B,p_2^B,p_2^G,k) = (1-\delta)\left[k\left(\hat{t}-p_2^G\right)+(1-k)\left(\bar{t}-p_2^G\right)\right]p_2^G - F \tag{2.15}$$

$$\text{with } \hat{t} := \frac{p_2^B-p_2^G}{\theta-1}$$

The sequence of events is the following:

- Period 1 is the period during which the brand-name drug is protected by a patent. It can be divided into two stages:

  - In stage 1, the brand-name firm chooses the optimal detailing level k_1.
  - In stage 2, the brand-name firm sets the monopoly price level.

- In period 2, the patent is expired and the brand-name firm faces potential competition from a generic firm. This period can be divided into three stages:
  - In stage 3, the generic firm decides about market entry.
  - In stage 4, the brand-name firm can increase the investment in advertising by k_2.
  - In stage 5, the brand-name firm and the generic firm set their optimal prices simultaneously, if there was generic market entry. Without market entry, the incumbent sets the monopoly price level.

In what follows, the game is solved by backward induction.

2.3 No Price Regulation

2.3.1 The Equilibrium of the Game

Stage 5: The Optimal Off-Patent Prices

After patent expiry, the entrant decides whether to enter the market or not. If there is no market entry, the monopoly price is set by the incumbent:

$$p_2^{B*}(k) = \frac{\bar{t}\theta}{2(k+\theta-\theta k)} \tag{2.16}$$

$p_2^{B*}(k)$ is found by maximising the monopoly profit in period 2 which can easily be derived from (2.13).

Given market entry, the optimal off-patent prices must form a Nash Equilibrium. The analysis focuses on an equilibrium in which $p_2^B > p_2^G$. Appendix 6.1.1 contains the exact conditions for this to be the only equilibrium. A sufficient condition is

$$\theta \geq \frac{1}{2k}\left[-k^2+2+\sqrt{k^4-5k^3+9k^2-8k+4}\right] \tag{2.17}$$

i.e. both k and θ must be sufficiently high.[8]

The generic firm sets a lower price than the brand-name firm. A higher generic price than the brand-name price would lead to zero profits for the entrant. Both detailed and not-detailed physicians would prescribe the brand-name version. Only if the generic price is lower than the brand-name price, the physicians are inclined to change their prescription behaviour towards the generic drug. The brand-name firm sets a higher price than the generic firm. On the one side, the brand-name firm foregoes the profits in the not-detailed market, however, on the other side, the optimal price in the detailed market is not distorted.

The optimal prices $p_2^{B*}(k)$ and $p_2^{G*}(k)$ are set simultaneously and are found by maximising the incumbent's and the entrant's second-period return functions.[9]

$$R_2^B(k) = (1-\delta)k\left[\bar{t}-\hat{t}\right]p_2^B \tag{2.18}$$

$$R_2^G(k) = (1-\delta)\left[k\left(\hat{t}-p_2^G\right)+(1-k)\left(\bar{t}-p_2^G\right)\right]p_2^G \tag{2.19}$$

$$\Rightarrow \quad p_2^{B*}(k) = \frac{\bar{t}(\theta-1)(k+2\theta-1)}{3k+4\theta-4} \tag{2.20}$$

$$\Rightarrow \quad p_2^{G*}(k) = \frac{\bar{t}(\theta-1)(2-k)}{3k+4\theta-4} \tag{2.21}$$

The incumbent's return depends only on the revenue in the detailed market, in which B serves the high-valuation patients. The price $p_2^{B*}(k)$ is set such that the marginal revenue on the detailed market is exactly zero.

The entrant serves both the treated low-valuation patients of the detailed physicians and all treated patients of the not-detailed physicians who prescribe the generic product to all types of patients whose valuation is at least as high as the price, because the price is lower and the perceived quality is the same. G also sets its price $p_2^{G*}(k)$ such that the overall marginal revenue is zero. However, since price differentiation is not possible, the marginal revenue on both the detailed and the not-detailed market is not equal to zero. The entrant would like to set a higher price in the not-detailed market and a lower price in the detailed market:

[8] Since k is an endogenous variable, whose optimal value is determined ex ante, it must be checked whether the profit-maximising k^* fulfils this condition. For the numerical illustration presented in section 2.3.3, this is the case.

[9] Since the advertising level has already been chosen in this stage and the fixed generic market entry costs are already sunk, these costs are irrelevant for the optimisation problem. Therefore, it is sufficient to maximise the return functions R_2^i.

$p_2^{G*}(\text{DM}) = \frac{1}{2}\bar{t}\frac{\theta-1}{2\theta-1}$ and $p_2^{G*}(\text{NDM}) = \frac{1}{2}\bar{t}$. The generic firm does not face competition in the not-detailed market. Thus, after a price increase, it loses some customers whose willingness to pay is lower than the price, but it does not lose any customers to the competitor. This competition effect exists in the detailed market, where a price increase additionally leads to some customers switching to the brand-name firm.

Proposition 1

(i) In equilibrium, the incumbent sets a higher price than the entrant: $p_2^{B*}(k) > p_2^{G*}(k)$.

(ii) $p_2^{B*}(k)$ *and* $p_2^{G*}(k)$ *decrease with the advertising level. The effect of advertising on the generic price is larger than on the brand-name price.*

Proof: (i) See the Appendix (6.1.1).

(ii)
$$\frac{dp_2^{B*}(k)}{dk} = -\frac{\bar{t}(\theta-1)(2\theta+1)}{(3k+4\theta-4)^2}$$
$$\frac{dp_2^{G*}(k)}{dk} = -\frac{2\bar{t}(\theta-1)(2\theta+1)}{(3k+4\theta-4)^2}$$
$$\frac{\bar{t}(\theta-1)(2\theta+1)}{(3k+4\theta-4)^2} < \frac{2\bar{t}(\theta-1)(2\theta+1)}{(3k+4\theta-4)^2}$$

If the advertising level increases, then the detailed market values more. This has no direct effect on the optimal brand-name price, because $p_2^{B*}(k)$ is already found by optimising only the return on the detailed market. But there is a direct effect on the optimal generic price. $p_2^{G*}(k)$ trades off the detailed and the not-detailed market. If k increases, then the detailed market weights more, where G would optimally set a lower price. If the generic price decreases, then this has an indirect effect on $p_2^{B*}(k)$ which decreases due to the competition effect. This reinforces further the decrease in the generic price as an indirect effect.

Given the above derived optimal prices in equilibrium, the off-patent profits can be expressed as

$$\pi_2^{B*}(k) = \frac{(1-\delta)\bar{t}^2 k(\theta-1)(k+2\theta-1)^2}{(3k+4\theta-4)^2} - \frac{1}{\gamma+1}k_2^{\gamma+1} \tag{2.22}$$

$$\pi_2^{G*}(k) = \frac{(1-\delta)\bar{t}^2(\theta-1)(2-k)^2(k+\theta-1)}{(3k+4\theta-4)^2} - F \tag{2.23}$$

Interestingly, the incumbent's off-patent return from advertising always increases with advertising:

$$\frac{dR_2^{B*}(k)}{dk} = \bar{t}^2(\theta-1)(k+2\theta-1)\frac{3k(k+2\theta-3)+4(\theta-1)(2\theta-1)}{(3k+4\theta-4)^3} \tag{2.24}$$

On the one hand, more advertising increases the detailed physicians' market share and, hence, the potential demand for the incumbent. On the other hand, prices decrease due to the fact that competition becomes more severe when more physicians are detailed. This price effect is dominated by the demand effect.[10]

Stage 4: The Optimal Level of Off-Patent Detailing

In stage 4, the incumbent decides on whether to increase the advertising level chosen during patent protection. He optimises his second-period profit with respect to k_2, the advertising level in period 2, given that he has already invested k_1 in period 1. The overall advertising level is composed of the individual advertising levels in both periods. k_2^* is defined by (2.26) with $k = k_1 + k_2^*$.

$$\pi_2^{B*}(k) = \bar{t}^2(\theta-1)(1-\delta)k\left(\frac{k+2\theta-1}{3k+4\theta-4}\right)^2 - \frac{1}{\gamma+1}k^{\gamma+1} \tag{2.25}$$

$$0 = \bar{t}^2(\theta-1)(1-\delta)(k+2\theta-1)\frac{3k(k+2\theta-3)+4(\theta-1)(2\theta-1)}{(3k+4\theta-4)^3} - k^{\gamma} \tag{2.26}$$

The optimal advertising level after generic market entry trades off the marginal off-patent return and the marginal costs of advertising.[11]

Stage 3: Market Entry Decision

Anticipating the optimal prices in stage 5 and thus the resulting profits, the entrant decides on market entry by comparing the profits with and without market entry. Market entry is profitable, if

$$\pi_2^{G*}(k) \geq 0 \quad \Leftrightarrow \quad \frac{\bar{t}^2(1-\delta)(\theta-1)(2-k)^2(k+\theta-1)}{(3k+4\theta-4)^2} - F \geq 0 \tag{2.27}$$

The decision depends on the (overall) level of advertising $k = k_1 + k_2$, which is determined in stage 1 and stage 4, on the fixed market entry costs F, the anticipated length of profitability $(1-\delta)$, and, finally, the perceived product differentiation θ.

Generic market entry is more likely for low generic market entry costs F, which is very intuitive, because the generic market entry costs have no indirect effects on advertising and

[10] Taking the equilibrium condition on θ determined in (2.17) into account, it can easily be verified that θ cannot simultaneously satisfy the necessary condition for (2.24) to become negative: $(k+2\theta-3) < 0 \iff \theta < \frac{1}{2}(3-k)$.

[11] The properties of k_2^* will not be investigated any further at this stage, because it will be shown in stage 1 that $k_2^* = 0$.

prices, but a direct negative effect on the generic profits.
The effects of both δ and θ on the anticipated generic profits are ambiguous, because there is both a direct and an indirect effect via the optimal prices and the optimal advertising level. It is not possible to calculate the effects analytically, therefore the discussion is deferred to the numerical illustration in section 2.3.3.

Advertising has an important effect on the market entry decision. Without any advertising ($k = 0$), there is only one (not-detailed) market in which all physicians perceive the two drugs as perfect substitutes. If G entered the market, then both firms would engage in fierce Bertrand competition. Hence, if the incumbent did not advertise during patent protection ($k_1 = 0$) and can credibly threat not to advertise after patent expiry ($k_2^* = 0$), then G anticipates that it could not cover the positive market entry costs and does not enter the market.

For any positive amount of advertising ($k > 0$), the generic profit can be positive. It is highest for $k \to 0$:

$$\pi_2^{G*}(k \to 0) = \frac{1}{4}(1-\delta)\bar{t}^2 - F \tag{2.28}$$

$$\frac{d\pi_2^{G*}(k)}{dk} = -\bar{t}^2(1-\delta)(2-k)\frac{3k\left[2(2\theta-1)+k\right]+4(2\theta-1)(\theta-1)}{(3k+4\theta-4)^3} < 0 \tag{2.29}$$

The more is invested in advertising, the smaller the generic profit, because the gain in the detailed market due to advertising is more than offset by the loss in the not-detailed market. This is due to the higher demand in the not-detailed market ($\bar{t} > \hat{t}$). Additionally, competition in the detailed market increases even further, such that G loses further demand in the detailed market. Overall, the effect of an increase in advertising is negative on the generic firm's profit and the generic firm would like the lowest positive advertising level possible.
The higher the advertising level, the more likely the generic returns are not sufficient to cover the positive generic market entry costs. This is the case for all advertising levels $k = k_1 + k_2$ above k_d, which is defined by

$$\frac{\bar{t}^2(1-\delta)(\theta-1)(2-k_d)^2(k_d+\theta-1)}{(3k_d+4\theta-4)^2} - F = 0 \tag{2.30}$$

This upper limit on advertising clearly depends negatively on the fixed market entry costs F.

Proposition 2 *Generic market entry is more likely for small generic market entry costs F and for a positive, but a small amount of advertising k.*

Stage 2: The Optimal On-Patent Price

The profit-maximising price $p_1^{B*}(k_1)$ during patent protection can be found by maximising the first-period return function of firm B. The resulting optimal price level depends on the ex ante chosen advertising level k_1 that will be derived afterwards.

$$R_1^B(k_1) = \delta\left[k_1\left(\bar{t}-\frac{1}{\theta}p_1^B\right)+(1-k_1)\left(\bar{t}-p_1^B\right)\right]p_1^B \tag{2.31}$$

$$\Rightarrow \quad p_1^{B*}(k_1) = \frac{\bar{t}\theta}{2(k_1+\theta-\theta k_1)} \tag{2.32}$$

The optimal price $p_1^{B*}(k_1)$ trades off the marginal return in the detailed market and in the not-detailed market. The more important the detailed market, the larger the price level, because it is more profitable for the incumbent to approach the optimal price of this market by further distorting the optimal price of the not-detailed market.

$$\frac{dp_1^{B*}(k_1)}{dk_1} = \frac{\bar{t}(\theta-1)\theta}{2(k_1+\theta-\theta k_1)^2} > 0 \tag{2.33}$$

Proposition 3 *The on-patent price level increases with advertising.*

Stage 1: The Optimal On-Patent Level of Detailing

The incumbent's profit during patent protection is given by

$$\pi_1^{B*}(k_1) = \frac{\delta\bar{t}^2\theta}{4(k_1+\theta-\theta k_1)} - \frac{1}{\gamma+1}k_1^{\gamma+1} \tag{2.34}$$

It is easy to verify that the return of advertising during the first period is positive. Since, in addition, the costs of advertising in both periods are assumed to be additive and convex, it is not optimal to delay any amount of advertising to the second period, but rather to only advertise during patent protection: $k^* = k_1^*$ and $k_2^* = 0$.

Proposition 4 *The brand-name firm advertises only during patent protection.*

This is in line with the empirical evidence in Caves et al. (1992) that advertising starts declining shortly before patent expiry and falls substantially with generic market entry.

Hence, the incumbent advertises only in the first stage. He anticipates that the threat to deter market entry by not investing in advertising is not credible, because as soon as there is generic market entry, advertising increases the incumbent's profit:

$$\frac{d\pi_2^{B*}(k)}{dk} \mid_{k=0} = \frac{\delta \bar{t}^2 (2\theta - 1)^2}{16(\theta - 1)} > 0 \tag{2.35}$$

Therefore, the incumbent can only deter generic market entry by over-investing in advertising. Whether this is profitable, is analysed in what follows.

If he accommodates market entry, then the incumbent takes into account that advertising has a positive effect on the on-patent price and a negative effect on the off-patent prices. k^* is found by maximising the overall profit $\pi_A^{B*}(k)$ and is defined by (2.37).

$$\pi_A^{B*}(k) = \delta \frac{\bar{t}^2 \theta}{4(k + \theta - \theta k)} + (1 - \delta)\bar{t}^2 (\theta - 1) k \left(\frac{k + 2\theta - 1}{3k + 4\theta - 4} \right)^2 - \frac{1}{\gamma + 1} k^{\gamma + 1} \tag{2.36}$$

$$(k^*)^\gamma = \bar{t}^2 (\theta - 1) \left[\frac{\delta \theta}{4\left(k^* + \theta - \theta k^*\right)^2} + (1 - \delta)(k^* + 2\theta - 1) \frac{3k^*(k^* + 2\theta - 3) + 4(\theta - 1)(2\theta - 1)}{(3k^* + 4\theta - 4)^3} \right] \tag{2.37}$$

The optimal level k^* exactly equalises the marginal return in both periods and the marginal cost of advertising.

The incumbent knows that there is no market entry for a sufficiently large stock of goodwill. Thus, if the incumbent decides to deter generic market entry, he needs to choose at least k_d, defined in (2.30), at which his overall profit is

$$\pi_D^{B*}(k_d) = \frac{\bar{t}^2 \theta}{4(k_d + \theta - \theta k_d)} - \frac{1}{\gamma + 1} k_d^{\gamma + 1} \tag{2.38}$$

The optimal brand-name price in this case is the monopoly price level, since no competition needs to be taken into account. Given this advertising level, the generic firm will not enter the market, because it anticipates the incumbent's price adjustment and the resulting generic losses due to the market entry costs.

Finally, the advertising level k_b might be optimal which blockades generic market entry without the need to change the profit-maximising behaviour. k_b is found by maximising the brand-name profit function without competition and is defined by (2.40).

$$\pi_B^{B*}(k_b) = \frac{\bar{t}^2 \theta}{4(k_b + \theta - \theta k_b)} - \frac{1}{\gamma + 1} k_b^{\gamma + 1} \tag{2.39}$$

$$k_b^\gamma = \frac{\bar{t}^2 \theta (\theta - 1)}{4(k_b + \theta - \theta k_b)^2} \tag{2.40}$$

A profit comparison between $\pi_A^{B*}(k^*)$, $\pi_D^{B*}(k_d)$, and $\pi_B^{B*}(k_b)$ helps the incumbent to decide whether to accommodate entry by choosing k^*, to deter entry by choosing k_d, or whether he can even blockade market entry by choosing k_b. The choice clearly depends on the parameter values and, the less lucrative the market for the potential generic entrant, the more likely the incumbent deters or even blockades market entry. This decision will be investigated further in the welfare analysis.

2.3.2 Welfare Analysis

Detailing targeted at physicians is in general allowed on the grounds that the informational advantages dominate the persuasive disadvantages (e.g. Scherer (2000), Leffler (1980)). Based on this argumentation, persuasive advertising does not provide any advantages, but is clearly wasteful by inducing brand-loyalty. The analysis above showed, however, that this might not be true, since even purely persuasive advertising can induce market entry.
Thus, in this context, it is interesting whether advertising can be beneficial from a social point of view, even though it might only be persuasive and not offer any informational content. Advertising can increase welfare, because it induces generic market entry and leads to lower prices after patent expiry and therefore to more patients being treated. Additionally, there is a market expanding effect on the detailed market due to the increased valuation.[12] The downside of advertising is, however, that on the one hand, it increases on-patent prices and it incorporates wasteful advertising costs, and on the other hand, over-investment in advertising can deter generic market entry.

The correct welfare analysis in the presence of (persuasive) advertising is a difficult problem. The question is whether to take the standard demand function or the demand function with the artificially increased valuation as basis to undertake the welfare analysis.[13] In the present model, only the physicians have different quality perceptions, whereas the patients

[12]It can be shown that the positive welfare aspect of advertising does not crucially depend on the distorted demand in the first period. However, if it were left out of the welfare function, the conditions on the parameters for advertising to be socially beneficial become stricter. The argument for the inclusion of the distorted demand in the welfare function is that, as long as prices are above marginal costs (here: $c = 0$), any demand increase is beneficial independent of whether the increase is based on the real valuation or on the distorted demand.

[13]See the discussion in Dixit and Norman (1978).

only care about the success of the medication. Since both the generic and the brand-name version are bioequivalent by assumption and should therefore be equally successful, welfare evaluations are based on the patients' actual valuation and not on the physicians' perceived valuation. This has the effect that advertising does not create any artificial welfare increase, because the quality is perceived to be higher.

Welfare in this setting is derived as follows: Both the pharmaceutical firms' profits and the patients' rent are taken into account. The profits and the rent are equally weighted. Since marginal costs were assumed to be zero, welfare is higher, the more patients are treated and thus the lower the prices. Given this definition of welfare, welfare with generic market entry and without generic market entry are, respectively:

$$\begin{aligned} W_A &= \delta\left[k^*\int_{\frac{1}{\theta}p_1^{B*}}^{\bar{t}} t\,dt + (1-k^*)\int_{p_1^{B*}}^{\bar{t}} t\,dt\right] + (1-\delta)\int_{p_2^{G*}}^{\bar{t}} t\,dt - \frac{1}{\gamma+1}(k^*)^{\gamma+1} - F \\ &= \frac{1}{2}\bar{t}^2\left[1-\delta\left(\frac{\theta}{2(k^*+\theta-\theta k^*)}\right)^2\left(k^*\frac{1}{\theta^2}+(1-k^*)\right) - (1-\delta)\left(\frac{(\theta-1)(2-k^*)}{3k^*+4\theta-4}\right)^2\right] - \\ &\quad -\frac{1}{\gamma+1}(k^*)^{\gamma+1} - F \qquad (2.41) \\ W_R &= k_r\int_{\frac{1}{\theta}p_r^{B*}}^{\bar{t}} t\,dt + (1-k_r)\int_{p_r^{B*}}^{\bar{t}} t\,dt - \frac{1}{\gamma+1}(k_r)^{\gamma+1} \\ &= \frac{1}{2}\bar{t}^2\left[1-\left(\frac{\theta}{2(k_r+\theta-\theta k_r)}\right)^2\left(k_r\frac{1}{\theta^2}+(1-k_r)\right)\right] - \frac{1}{\gamma+1}(k_r)^{\gamma+1}, \quad R=D,B,\ r=d,b \\ & \qquad (2.42) \end{aligned}$$

In both cases, welfare considers, how many patients are treated in the two periods. However, the price levels differ: Without generic market entry, the brand-name firm sets the monopoly price p_r^{B*} in both periods. Due to the distorted quality perception on the detailed market, demand is higher than on the not-detailed market. The indifferent patients are determined by $t = \frac{1}{\theta}p_r^{B*}$ respectively $t = p_r^{B*}$.

With generic market entry, the situation differs during patent protection and after patent expiry. The relevant off-patent price is p_2^{G*}, because this is the lower price level of the two and determines the patient that is just indifferent between being treated or not. The generic drug is prescribed by either a detailed or a not-detailed physician. Both attach the valuation $\theta = 1$ to the generic version such that the total amount of treated patients is not distorted by advertising. During patent protection, the same demand distortion applies as without

parameter	k^*	k_d	k_b	$\pi_A^B(k^*)$	$\pi_D^B(k_d)$	$\pi_B^B(k_b)$	$\pi_2^G(k^*)$	W
$\theta = 1.7$	0.753	0.931	0.779	0.248	0.319	/	D	0.285
$\theta = 2.0$	0.810	0.991	0.850	0.306	0.361	/	D	0.240
$\theta = 2.2$	0.842	1.017	0.898	0.345	/	/	0.012	0.364
$\theta = 2.4$	0.873	1.037	0.949	0.386	/	/	0.011	0.351
$\delta = 0.2$	0.827	1.134	0.898	0.321	/	/	0.025	0.394
$\delta = 0.4$	0.842	1.017	0.898	0.345	/	/	0.012	0.364
$\delta = 0.5$	0.851	0.937	0.898	0.358	0.421	/	D	0.275
$\delta = 0.6$	0.859	0.832	0.898	0.370	0.418	0.423	B	0.295
$\gamma = 5$	0.810	1.017	0.868	0.330	/	/	0.014	0.359
$\gamma = 6$	0.842	1.017	0.898	0.345	/	/	0.012	0.364
$\gamma = 7$	0.866	1.017	0.917	0.357	/	/	0.010	0.369
$\gamma = 10$	0.907	1.017	0.948	0.381	/	/	0.007	0.380

Table 2.1: A Numerical Illustration.

generic market entry, but with respect to the optimal price p_1^{B*}.
Additionally, the advertising costs are considered and, if appropriate, the generic market entry costs. If generic market entry is deterred, then over-investment in advertising leads to higher advertising costs than the situation with accommodated entry.

2.3.3 A Numerical Illustration

Unfortunately, it cannot be solved for any of the k explicitly, therefore the solution has to be approximated numerically. The default values are $\bar{t} = 1$, $\delta = 0.4$, $\theta = 2.2$, $\gamma = 6$, and $F = 0.025$. Some numerical comparative statics is undertaken for those parameters, that are thought of as most easily manipulated by the regulator. These are the detailing costs γ, the patent length δ, and the efficiency of advertising θ. Since the effects due to the entry costs F have already been discussed in section 2.3.1, they will not be illustrated numerically.

For the interpretation, it is important, how changes in the parameter values affect the entrant's expected profit, because this determines whether entry is deterred, blockaded, or accommodated. Furthermore, the effect on the profitability of advertising is relevant, because this determines the optimal level of detailing. These effects are responsible for the resulting welfare. Table 2.1 summarises the numerical results.

- θ has a direct and an indirect effect on generic market entry. Since it represents the degree of vertical product differentiation, competition is less severe, the larger is θ. However, a high θ increases the investment in advertising, which reduces the generic profit as an indirect effect. In this example, a small θ, i.e. a rather ineffective advertising technology, makes it more likely that generic market entry is deterred, because the two drug versions are only slightly differentiated and, thus, competition is strong. Only for a sufficiently high product differentiation, market entry cannot be deterred. Welfare is higher with generic market entry due to the additional lower-priced drug which enables more patients to be treated. As long as generic market entry is guaranteed, θ should be as low as possible for welfare to be maximised. The reason is that advertising k increases with θ which increases the advertising costs.

- The smaller δ, i.e. the period of patent protection, the more likely generic market entry is accommodated. $(1 - \delta)$ represents the expected length of the period during which the generic version can be sold profitably. This period is the shorter, the longer the original drug is protected by a patent. Additionally, a small δ reduces the investment in k which increases the expected generic profit also indirectly. For the incumbent, on the contrary, the profit is the higher, the larger δ, because the period is longer during which he can act as monopolist. As expected, welfare is maximised for a short period of patent protection such that generic substitutes are offered and the prices are lower for a longer period.

- The advertising costs γ do not directly affect the probability of generic market entry, because the entrant's profit is independent of γ. Although a high γ, i.e. low advertising costs, increases k^*, welfare increases for low advertising costs due to the direct cost effect as long as generic market entry is not deterred.

- Unsurprisingly, welfare increases with low market entry costs F. They directly affect the probability of generic market entry and they directly reduce welfare.

Although only a few examples are presented, the basic intuition can be derived.[14] Welfare is higher, if a generic substitute is offered after patent expiry. This is more likely, if the drug versions are sufficiently differentiated, if the period of patent protection is rather short, and

[14] Additional numerical examples show that these results are robust to changes in the parameter values. They can be provided by the author upon request.

if the fixed market entry costs are low. The regulator cannot affect the incumbent's choice of the advertising level and, hence, the effect on the generic firm's market entry decision. But he has the possibility to manipulate the conditions under which the incumbent chooses the profit-maximising advertising level. This will be further discussed in the following section.

2.3.4 Policy Implications

Product differentiation between the original drugs and their generic substitutes is necessary in order to induce competition and therefore lower off-patent prices. This differentiation must be sufficiently high in order to avoid that generic market entry is deterred by over-investing in detailing. Furthermore, the analysis showed that, given generic market entry, welfare is higher, the lower this differentiation is. The degree of product differentiation is basically determined by the effectiveness of advertising, i.e. how persuasive the pharmaceutical firms' detailing representatives are. Certainly, the regulator cannot directly influence them, but the effectiveness can be influenced indirectly. The health authority can provide information about the therapeutical equivalence between brand-name and generic drugs. Furthermore, the extent to which physicians are allowed to receive gifts or sponsored (conference) trips can be regulated in order to reduce the felt obligation towards specific brands. This might be a way to reduce the effect that detailing has on the individual physician.

The optimal *patent length* is set such that the pharmaceutical firm is granted sufficient time of monopoly power during which it can recoup the huge sunk innovation costs. But taking into account that advertising expenditures easily even exceed R&D-expenditures, the effect of patent protection on advertising and generic market entry should not be neglected, when deciding about the optimal patent length. As Taggart (1993) has pointed out, however, the effective patent length is already rather short and thus seems to work in favour of generic market entry.[15]

First intuition suggests that persuasive advertising should be minimised by making detailing harder, i.e. high *advertising costs* are socially beneficial. But it has been shown in the

[15]The nominal patent life is 20 years in the European Union and 17 years in the United States, but because the patent life includes some of the initial time required for clinical testing and approval, the patent life effectively reduces to e.g. 6.4 years in Germany, 8.7 years in the United Kingdom, and 9.7 years in the United States (see Taggart (1993)).

analysis that low advertising costs can be preferable, because they reduce the direct welfare loss. Hence, the health authorities should not try to make detailing absolutely harder, e.g. by imposing more bureaucracy, but rather to reduce these costs of detailing.

Any measures that reduce the generic *market entry costs* increase welfare. The direct setup-costs cannot be affected by the health authorities, but by reducing the bureaucracy and formal requirements associated with market entry, the setup-costs can be affected indirectly. A step in the right direction was the introduction of the Waxman-Hatch Act in 1984, which reduced the generic firms' financial burden to replicate the original drug's tests. The only requirement now is the demonstration that the generic substitute has the same active ingredients as the original version, that the absorbation rate is similar, and that it is manufactured properly (Scherer (2000)). The Waxman-Hatch Act increased welfare both directly by reducing the welfare costs, and indirectly by spurring further generic market entry.

2.4 Price Regulation

An interesting extension of the model applies to price regulation. Until now, the model assumes that there is free pricing in the pharmaceutical market, which is basically only true for the United States. In the European market, most countries restrict pharmaceutical prices in some way. This is likely to affect the equilibrium outcome.

There is empirical evidence that generic market shares are lower in countries where price regulation is stricter (Danzon and Chao (2000)). Countries that allow for (relatively) free pricing in the pharmaceutical market, as certainly the United States, but also the United Kingdom or Germany, have higher generic market shares than for example Italy or France, where both countries have strict price regulations. The aim of this extension of the model is to give an explanation for the effect of price regulation on generic market entry, emphasising especially the possibility of brand-name advertising. It will be shown that, within this framework, generic market entry is the more likely deterred, the stricter price regulation is.

The framework must be adjusted only marginally, because the optimal prices in both periods might be constrained. Out of the many variants of price regulations, e.g. price caps, reference pricing, or the indirect form of price regulation via profit limits, the simplest version is assumed. A regulator sets a price cap $\bar{p}$ that cannot be surpassed.

2.4.1 The Equilibrium of the Game

In order to derive the equilibrium of the game, four different cases need to be considered, which depend on the level of the price cap and the brand-name firm's pricing strategy over time:

1. The price cap is never binding.
2. The on-patent brand-name price is constrained, whereas the off-patent one is unconstrained.
3. The off-patent brand-name price is constrained, but not the on-patent one.
4. The brand-name price is constrained in both periods.

Empirical studies show that the original brand-name drug's price can both decrease or increase after generic market entry (see Frank and Salkever (1997)). This observation can be replicated by models of vertical product differentiation, like the present one. The intuition is that advertising creates market segmentation. After patent expiry, the original drug focuses on the more price-inelastic market segment, whereas it was sold to both segments during patent protection. This enables the brand-name firm to increase its optimal price after patent expiry under certain circumstances (in this model for a high θ and a low k), although it faces generic competition.

In order to limit the number of cases that might possibly arise to the most relevant one in practice, the price cap is assumed to be sufficiently low to restrict the optimal brand-name price both during patent protection and after patent expiry. This means that only case (4) is analysed.

Stage 5: The Optimal Off-Patent Prices

If there was market entry in stage 3, then there exists an equilibrium in pure strategies in which the incumbent simply sets the highest price level possible, i.e. the price cap $\bar{p}$, to which the entrant reacts which a lower generic price.[16] This requires some restrictions on

[16]If there was no market entry, the brand-name firm sets the highest possible price, $p_2^{B*} = \bar{p}$, since the price ceiling is binding by assumption.

the parameters which are derived in the Appendix (6.1.2). They can be summarised by $\bar{p} \in (p_{low}, p_{high})$, i.e. the price cap is bounded from below and above.

The price cap is binding and the incumbent would like to set a higher price. Therefore, he chooses the highest possible price level to which the generic firm reacts with a lower price in order to guarantee that the generic substitute is prescribed to the low-valuation patients in the detailed market and is used as only treatment in the not-detailed market. The optimal generic response to $p_2^{B*} = \bar{p}$ can be found by maximising the entrant's second-period profit function:

$$\pi_2^G(k) = (1-\delta)\left[k\left(\frac{\bar{p}-p_2^G}{\theta-1}-p_2^G\right)+(1-k)\left(\bar{t}-p_2^G\right)\right]p_2^G \tag{2.43}$$

$$\Rightarrow \quad p_2^{G*}(k) = \frac{\bar{t}(\theta-1)(1-k)+k\bar{p}}{2(k+\theta-1)} \tag{2.44}$$

The optimal price level $p_2^{G*}(k)$ trades off the detailed and the not-detailed market. If price discrimination was possible, G would like to set a higher price in the not-detailed market than in the detailed market:

$$p_2^{G*}(DM) = \frac{\bar{p}}{2\theta}$$

$$p_2^{G*}(NDM) = \begin{cases} \frac{1}{2}\bar{t} & \text{if} \quad \frac{1}{2}\bar{t} \leq \bar{p} \\ \bar{p}-\epsilon, \quad \epsilon \to 0 & \text{if} \quad \frac{1}{2}\bar{t} > \bar{p} \end{cases}$$

G does not face any competitor in the not-detailed market and does not lose patients to the brand-name firm after a price increase. This competition effect exists, however, in the detailed market.

Proposition 5

(i) In equilibrium, the incumbent sets the maximally allowed price cap to which the entrant optimally reacts with a lower price level.

(ii) The generic price level is a strategic complement to the price cap.

(iii) The generic price level decreases, the more is invested in advertising.

Proof: (i) See the Appendix (6.1.2).

(ii) $\frac{dp_2^{G*}(\bar{p})}{d\bar{p}} = \frac{k}{2(k+\theta-1)} > 0$

(iii) $\frac{dp_2^{G*}(k)}{dk} = -\frac{(\theta-1)(\bar{t}\theta-\bar{p})}{2(k+\theta-1)^2} < 0$

The off-patent prices are strategic complements. If one firm decreases its price level, then more patients switch to this firm. In order to countervail this effect, the other firm must also decrease its price. A lower price cap is equivalent to a lower brand-name price level. Thus, a lower price cap results in a lower generic price $p_2^{G*}(k)$.
The optimal generic price trades off the detailed and the not-detailed market, and G would optimally set a lower price in the detailed market with branded competition. If k increases, then the detailed market weights more and the optimal price $p_2^{G*}(k)$ decreases.

Stage 4: The Optimal Off-Patent Level of Detailing

Like in the situation without price regulation, the incumbent can decide to increase his on-patent advertising level in stage 4. However, with the same argumentation as before, it can be shown that it is optimal to advertise only during patent protection, because on the one hand, the marginal costs remain the same, and on the other hand, this enables the incumbent to raise the benefits from detailing in *both* periods.

Stage 3: The Market Entry Decision

In stage 3, the generic firm decides on whether to enter the market or not. The generic market entry decision depends on G's expected profits:

$$\pi_2^{G*}(k) = \frac{(1-\delta)\left[\bar{t}(\theta-1)(1-k)+k\bar{p}\right]^2}{4(\theta-1)(k+\theta-1)} - F \tag{2.45}$$

Clearly, there is no market entry, if there is no advertising and hence no product differentiation, because fierce Bertrand competition would result in zero returns which are not sufficient to cover the sunk market entry costs. Hence, if the incumbent did not advertise during patent protection ($k_1 = 0$) and can credibly threat not to advertise after patent expiry ($k_2^* = 0$), then there is no generic market entry.

Furthermore, it can be shown that the generic profits decrease, the stricter price regulation is, and the more is invested in brand-name advertising:

$$\frac{d\pi_2^{G*}(k)}{d\bar{p}} = \frac{(1-\delta)k\left[\bar{t}(\theta-1)(1-k)+k\bar{p}\right]}{2(\theta-1)(k+\theta-1)} > 0 \tag{2.46}$$

$$\frac{d\pi_2^{G*}(k)}{dk} = \frac{(1-\delta)\left[\bar{t}(\theta-1)(1-k)+k\bar{p}\right]\left[\bar{t}(\theta-1)(1-k)+k\bar{p}-2(\theta-1)(\bar{t}\theta-\bar{p})\right]}{4(\theta-1)(k+\theta-1)^2} \tag{2.47}$$

$$< 0 \Leftrightarrow \bar{p} < \bar{t}(\theta-1)\frac{2\theta-1+k}{k+2\theta-2} \tag{2.48}$$

$\bar{p}$ fulfils the condition in (2.48) as long as the price cap is binding in both periods.[17]

Proposition 6 *The lower the price cap, the more likely generic market entry is deterred.*

This result is in line with the empirical evidence in Danzon and Chao (2000) who find a negative relationship between price regulation and generic competition. The model confirms that the stricter price regulation, the less profitable the pharmaceutical market within the specific country, and the less likely a generic producer enters this market.

The more the brand-name firm invests in advertising and thus the more the brand-name firm invests in its stock of goodwill, the lower is the generic profit. k_d is the lower bound on advertising that just deters market entry:
$\pi_2^{G*}(k) < 0 \quad \Leftrightarrow \quad k > k_d$

$$k_d := \frac{(\theta-1)\cdot\left[2F + \bar{t}(1-\delta)\left(\bar{t}(\theta-1)-\bar{p}\right) - 2\sqrt{F\left[F+(1-\delta)\left(\bar{t}(\theta-1)-\bar{p}\right)\left(\bar{t}\theta-\bar{p}\right)\right]}\right]}{(1-\delta)\left(\bar{t}(\theta-1)-\bar{p}\right)^2} \tag{2.49}$$

This upper limit of advertising, that just guarantees market entry, is lower than without price regulation, i.e. $k_d > k_d(PR)$. This can be verified by analysing the fixed market entry costs which just guarantee generic market entry. For any advertising level, they must be lower with price regulation ($\bar{F}(PR)$) than without market entry ($\bar{F}$):

$$\bar{F} \quad \leq \quad \frac{\bar{t}^2(1-\delta)(\theta-1)(2-k)^2(k+\theta-1)}{(3k+4\theta-4)^2} \tag{2.50}$$

$$\bar{F}(PR) \quad \leq \quad \frac{(1-\delta)\left[\bar{t}(\theta-1)(1-k)+k\bar{p}\right]^2}{4(\theta-1)(k+\theta-1)} \tag{2.51}$$

$\bar{F}(PR) \leq \bar{F}$ for all $0 \leq \bar{p} \leq p_2^{B*}$.

Proposition 7 *With a binding price cap, generic market entry is more likely deterred by over-investing in advertising as compared to the situation without price regulation.*

Stage 2: The Optimal On-Patent Price

The optimal brand-name price during patent protection is higher than the price cap by assumption. Therefore, the incumbent sets the highest possible price level $\bar{p}$ in stage 2.

[17]The unconstrained brand-name prices are $p_1^{B*}(k) = \frac{\bar{t}\theta}{2(k+\theta-\theta k)}$ and $p_2^{B*}(k) = \frac{\bar{t}(\theta-1)(k+2\theta-1)}{3k+4\theta-4}$. The derivation can be found in section 2.3.1. The condition in (2.48) is always fulfilled, given that $\bar{p} \leq p_1^{B*}, p_2^{B*}$.

Stage 1: The Optimal Level of Detailing

In stage 1, finally, the incumbent decides how much to invest in advertising by anticipating the effect of advertising on his overall profit and on the generic firm's entry decision in stage 3. The incumbent knows that there is no market entry for a sufficiently large stock of goodwill ($k > k_d$), and that he cannot deter market entry by not advertising in the first period at all, because advertising increases the incumbent's profit as soon as there is generic market entry. Furthermore, like in the situation without price regulation, it can be shown that it is not optimal to delay the investment in advertising to the second period, but rather to only advertise during patent protection: $k_2^* = 0$ and $k_1^* = k^*$.

If the incumbent chooses to accommodate market entry, then the overall brand-name profit is given by

$$\pi_A^B(k) = \delta\left[\bar{t} - \bar{p}\left(k\frac{1}{\theta} + 1 - k\right)\right]\bar{p} + (1-\delta)k\bar{p}\frac{(k+2\theta-1)(\bar{t}(\theta-1)-\bar{p})+\bar{p}}{2(\theta-1)(k+\theta-1)} - \frac{1}{\gamma+1}k^{\gamma+1} \tag{2.52}$$

In this case, the optimal advertising level k^* can be found by maximising the profit function $\pi_A^B(k)$, and is defined by the following first-order condition:

$$\delta\bar{p}^2\left(1-\tfrac{1}{\theta}\right) + \tfrac{(1-\delta)\bar{p}}{2(k^*+\theta-1)^2}\left[\bar{t}\left(\theta(2\theta-1)-(1-k^*)(2\theta-1+k^*)\right) - \bar{p}\left((\theta-1)+\tfrac{(\theta-1+k^*)^2}{(\theta-1}\right)\right] = (k^*)^\gamma \tag{2.53}$$

If the incumbent decides to deter generic market entry, then he needs to set at least k_d, defined in (2.49), at which his overall profit is

$$\pi_D^B(k_d) = \left[\bar{t} - \bar{p}\left(k_d\frac{1}{\theta} + 1 - k_d\right)\right]\bar{p} - \frac{1}{\gamma+1}(k_d)^{\gamma+1} \tag{2.54}$$

Finally, the advertising level k_b might be optimal that blockades generic market entry without the need to change the profit-maximising behaviour. The optimal advertising choice automatically blockades generic market entry. k_b is found by maximising the brand-name profit function without competition.

$$\pi_B^B(k_b) = \left[\bar{t} - \bar{p}\left(k_b\frac{1}{\theta} + 1 - k_b\right)\right]\bar{p} - \frac{1}{\gamma+1}k_b^{\gamma+1} \tag{2.55}$$

$$k_b = \exp\left\{\frac{1}{\gamma}\ln\left(\bar{p}^2\frac{\theta-1}{\theta}\right)\right\} \tag{2.56}$$

A profit comparison between $\pi_A^B(k^*)$, $\pi_D^B(k_d)$, and $\pi_B^B(k_b)$ helps the incumbent to decide whether to accommodate entry by choosing k^*, to deter entry by choosing k_d, or whether

$\bar{p}$	k^*	k_d	k_b	$\pi_A^B(k^*)$	$\pi_D^B(k_d)$	W
0.50	0.852	1.035	0.812	0.287	/	0.442
0.49	0.851	1.019	0.809	0.284	/	0.443
0.48	0.850	1.003	0.806	0.281	/	0.444
0.47	0.849	0.988	0.802	0.277	0.279	0.392
0.46	0.848	0.973	0.799	0.274	0.284	0.404

Table 2.2: Numerical Example

he can even blockade market entry by choosing k_b. Since the generic profit $\pi_2^G(k^*)$ decreases for lower price caps $\bar{p}$, all else equal, as it was derived in (2.46), it becomes more likely that market entry is deterred, the stricter price regulation is.

Unfortunately, the profits cannot be compared analytically, because k^* cannot be expressed explicitly. Therefore, a numerical example is presented in Table 2.2 as an illustration, how the incumbent's decision depends on the price cap.[18] Although I restrict myself to only one example, the basic intuition can nicely be derived. The stricter price regulation is, i.e. the lower the price cap is, the less the incumbent invests in the optimal market entry-accommodating advertising k^*. This is intuitive, because the incumbent is restricted in his price level and cannot optimally skim the artificially increased quality perception θ.
Whereas generic market entry can never be blockaded in this example ($k_b < k_d$), it can easily be seen that market entry is more likely deterred, the lower the price cap is. This is due to the direct negative effect of a lower price cap on generic profits. In order to guarantee sufficiently high profits even in the presence of strict price regulation, the advertising level must be sufficiently low to counteract the negative effect of the low price ceiling.

In this specific example, for any price cap $\bar{p} \geq 0.48$, the market entry-deterring advertising level is too large with $k_d > 1$. The incumbent maximises his profits by choosing k^* and by accommodating market entry. For a lower price cap, the brand-name profits are higher, when market entry is deterred.

[18]The following parameter values are used: $\theta = 2, \bar{t} = 1, \delta = 0.3, \gamma = 10$, and $F = 0.02$. The unconstrained brand-name prices for these parameter values are $p_1^{B*} = 0.892$ and $p_2^{B*} = 0.584$. These values present the upper limit on a price cap that is binding in both periods. This is only an illustrative example, but the results do not change qualitatively, when other parameter values are used.

Hence, generic market entry is more likely, the higher the price cap, i.e. the more lenient the price regulation scheme. The stricter the price regulation scheme, the more likely market entry is deterred or even blockaded.

2.4.2 Welfare Implications

Price regulation increases the danger that generic market entry is deterred (or even blockaded) by the incumbent by over-investing in pre-entry advertising, i.e. by creating a large stock of goodwill. This is more likely the case, the lower the price cap is. Hence, there is a trade-off with respect to welfare in this static setting: On the one hand, a low price cap reduces prices and increases simultaneously (patients') welfare. On the other hand, strict price regulation makes it more likely that generic market entry is deterred and that there is no lower-priced drug version. Hence, in the welfare analysis, welfare with generic market entry, W_A, and welfare without generic market entry, W_D respectively W_B, need to be compared. The definition of welfare is the same as in the analysis without price regulation.

$$\begin{aligned} W_A &= \delta\left[k^*\int_{\frac{1}{\theta}\bar{p}}^{\bar{t}} t\,dt + (1-k^*)\int_{\bar{p}}^{\bar{t}} t\,dt\right] + (1-\delta)\int_{p_2^{G*}}^{\bar{t}} t\,dt - \frac{1}{\gamma+1}(k^*)^{\gamma+1} - F \\ &= \frac{1}{2}\left[\bar{t}^2 - \delta\bar{p}^2\left(\frac{k^*}{\theta^2} + (1-k^*)\right) - (1-\delta)\left(\frac{\bar{t}(\theta-1)(1-k^*) + k^*\bar{p}}{2(k^*+\theta-1)}\right)^2\right] - \frac{1}{\gamma+1}(k^*)^{\gamma+1} - F \end{aligned} \tag{2.57}$$

Welfare with generic market entry considers the advertising and the generic market entry costs and how many patients are treated in both periods. The relevant off-patent price is p_2^{G*}, because this is the lower price level of the two and determines the patient that is just indifferent between being treated or not. The generic drug is prescribed by either a detailed or a not-detailed physician. Both attach the valuation $\theta = 1$ to the generic version such that the total amount of treated patients is not distorted by advertising.
Without price regulation, allowing for advertising involves two countervailing effects on the welfare during patent protection. For given prices, demand increases due to the fact that the detailed physicians attach a higher value to the brand-name drug. However, advertising also increases on-patent prices, which in turn reduces demand. This second negative effect is irrelevant for a binding price cap. Due to the first effect, the indifferent patient in the not-detailed market is determined by $t = \bar{p}$, in the detailed market by $t = \frac{1}{\theta}\bar{p}$.

Welfare without generic market entry is derived analogously, but no generic market entry

costs enter the welfare function. The price level is the same in both periods and equal to the price cap $\bar{p}$. This situation arises, if the advertising level is sufficiently high: $k > k_d, k_b$.

$$\begin{aligned} W_R &= k_r \int_{\frac{1}{\theta}\bar{p}}^{\bar{t}} t \, dt + (1 - k_r) \int_{\bar{p}}^{\bar{t}} t \, dt - \frac{1}{\gamma + 1} k_r^{\gamma+1} \\ &= \frac{1}{2} \left[\bar{t}^2 - \bar{p}^2 \left(\frac{k_r}{\theta^2} + (1 - k_r) \right) \right] - \frac{1}{\gamma + 1} k_r^{\gamma+1}, \quad R = D, B, \quad r = d, b \end{aligned} \tag{2.58}$$

Again, no analytical solution is possible. Therefore, the same numerical illustration is used as before in order to show an example for which a strict price regulation regime is worse for welfare than a lenient price regulation regime, although it results in lower brand-name prices. The results can be seen in Table 2.2, where welfare at a higher price cap but with generic market entry accommodation is higher than with a lower price cap and no generic market entry. For a very strict price regulation, welfare can, of course, again become higher than with high price caps, although there is no generic market entry. Generic competition is in general beneficial for welfare, because it induces lower pharmaceutical prices. If the price regulation in a country is already so strict that it implements very low pharmaceutical prices, no generic competition is necessary. However, other considerations like the effect of a very low price cap on innovation incentives or on the brand-name firms' market entry decision might curtail, how low the price ceiling can be set in practice.

2.5 Conclusion

This chapter is concerned with a pharmaceutical market in which, after the expiry of a patent, the incumbent faces the threat of generic market entry. During patent life, the incumbent has the possibility to invest in advertising targeted at the prescribing physician with the effect that the brand-name drug's perceived quality is increased and differentiated from the generic drug's quality. On the one hand, this is an important precondition for generic market entry, on the other hand, generic market entry can be deterred by over-investing in advertising.

Detailing in the pharmaceutical market is in general allowed, because it is assumed that it is mostly informative. Hence, the positive effects of a better matching between patients and drugs dominate the negative effects attributed to persuasive advertising. The analysis showed, however, that there are some positive welfare aspects with respect to even persuasive advertising in the sense that generic market entry can be induced and post-patent prices are

lower. Hence, against first intuition, the persuasive effect of advertising does not necessarily decrease the welfare benefit of advertising due to information, but can even reinforce it under certain circumstances. This is particularly true for a short period of patent protection, for a sufficiently effective advertising technology, and low market entry costs. Hence, these are the parameters that a regulator should try to manipulate in order to ensure generic market entry and to maximise welfare.

Furthermore, the model is able to explain the empirical finding that strict price regulation can reduce generic competition. Price regulation affects and reduces both the brand-name and the generic off-patent prices. This in turn reduces the generic firm's expected return from market entry and makes it more likely that the incumbent can deter generic market entry by over-investing in pre-entry advertising, i.e. by over-investing in the stock of goodwill. This can reduce the generic firm's return by so much that it is no longer sufficient to cover the entry costs.

The model is of course severely simplified and important aspects have not been taken into account. The inclusion of a principal-agent relationship between the patients and their physicians, for example, might stress the downsides of advertising. Additionally, the aspect of competition has been limited to the brand-name producer and his first generic competitor. The effect of advertising on the total number of generic entrants would be an interesting extension.
Important aspects of advertising certainly also concern the competition between several brand-name companies. However, between different brand-name drugs, *informative* advertising seems to be more appropriate which gives the information about the presence of an additional, therapeutically equivalent, but horizontally differentiated (on-patent) drug. This chapter focuses on purely persuasive advertising. Therefore, this aspect has not been investigated any further.

Although a very basic form of price regulation is examined in the extension, the results should not change qualitatively without this simplification. The result is based on the effect that the generic profit after market entry is the lower, the lower the brand-name firm must set its price due to price regulation. This effect does not depend on the exact design of the price regulation regime. Hence, although price regulation in the European markets is varied and more complicated than presented in the paper, the results are applicable in practice.

Chapter 3

Reference Pricing

3.1 Introduction

The pharmaceutical market is characterised by a price-inelastic demand mainly due to extensive medical insurance. Since individuals, once they are ill, only pay a small fraction of their medical costs, prices are likely to have a limited effect. This does not only affect the choice of whether or not to consume a drug, but also the choice between alternative drug treatments. On the supply-side, there are large sunk R&D costs associated with the discovery of new drug treatments. In order to stimulate innovation, pharmaceutical firms are granted market power (for a given period) by patent protection.

The combination of supply-side market power and price-inelastic demand has induced purchasers to employ various means to control medical expenditures.[1] Basically, one can distinguish between two price control mechanisms: Direct and indirect regulation of drug prices, where the later regulates the reimbursement level and is frequently referred to as reference pricing. While direct price regulation limits the pharmaceutical firms' ability to exploit their market power by charging high prices, reference pricing aims at stimulating competition by making demand more price-elastic. In this chapter, the effects of reference pricing on the price-setting strategies of the pharmaceutical firms are analysed. On the basis of this analysis, the implications for market entry of new drug treatments, patient health risks, and

[1] Danzon (1997a) provides an excellent overview and discussion of various regulatory mechanisms in the pharmaceutical industry.

optimal drug reimbursement policies are discussed. While these issues have received some empirical attention, theoretical contributions are very limited. According to the extensive literature survey by Lopez-Casasnovas and Puig-Junoy (2001), the bulk of the reference pricing literature is mainly descriptive, and there is a pronounced lack of theoretical studies analysing the effects of reference price systems (see also Danzon (2001)).

Reference pricing of prescription drugs is quite novel, but it has rapidly become a widely used price control mechanism in the pharmaceutical market. Germany's Statutory Health Insurance System, generally viewed as the pioneer in this regard, introduced reference pricing for prescription drugs in 1989 and was followed in Europe by the Netherlands in 1991, Denmark and Sweden in 1993, Spain in 2000, and Belgium and Italy in 2001. Norway adopted reference pricing in 1993, but abandoned it in 2001, because the expected cost savings did not materialise. Outside Europe, reference pricing has been adopted by Australia, the Canadian province of British Columbia, and New Zealand.[2]

The reference price system is constructed as follows: Drugs are classified into clusters based on similar therapeutic effects. The regulator sets a reference price based on a relatively low-priced drug in the cluster, e.g. the minimum or median price. The reference price is the maximum reimbursement for all products in the group. Pharmaceutical firms can set their prices above the reference price, but, in this case, the patient must pay the surcharge.

The construction of therapeutic clusters for reference pricing is by far the most controversial task in the development of such systems. These clusters may be narrowly or broadly defined:

(i) products with the same active chemical ingredients,

(ii) products with chemically related active ingredients that are pharmacologically equivalent, and

(iii) products that may be neither chemically identical nor pharmacologically equivalent, but have comparable therapeutic effects.

[2]In the United States, reference pricing has been proposed as a possible approach to drug reimbursement for a comprehensive Medicare drug benefit (Huskamp et al. (2000)). Kanavos and Reinhardt (2003) argue that reference pricing for drugs is compatible with the United States' health care. Notably, generic reference pricing is well-established in the US through "maximum allowable charge" programs used by, e.g., Medicaid.

By its nature, the first type of cluster includes only off-patent brand-name drugs and their generic substitutes. The second and third may include on-patent drugs. They differ in breadth, but are qualitatively similar. The first type is referred to as generic reference pricing (GRP), and the second and third as therapeutic reference pricing (TRP).

The model in this chapter is constructed such that the effects of the two reference price systems, TRP and GRP, can be analysed, as well as the benchmark case of no reference pricing (NRP) where patients pay a fixed share (co-payment rate) of the drug price. The basic setup is a therapeutic market with potentially three pharmaceutical firms, where two of the firms offer original brand-name drugs with different chemical ingredients. One of the brand-name drugs is an old treatment, e.g. the breakthrough drug, that has lost its patent protection and faces competition from a third firm offering a generic version, which is perceived to be of lower quality than the off-patent brand-name drug.[3] The other brand-name drug is a new, horizontally differentiated treatment under patent protection that will be introduced in the market, if the profits are sufficient to cover the entry costs.[4] With this modelling approach, the arguments for and against reference price systems in general, and between TRP and GRP in particular can be discussed.

The main argument in favour of reference pricing is that it stimulates price competition by making demand more elastic and thus resulting in lower medical expenditures. Intuitively, the effect on price competition should be stronger, the wider the cluster is defined. The model confirms this line of argument. The price of every drug in the therapeutic market is highest under NRP and lowest under TRP. It is worth noting that GRP not only reduces the prices of the drugs in the reference cluster, but also puts a downward pressure on the price of the not-included, but therapeutically equivalent drug. This is due to the prices being strategic complements.[5]

[3]Empirical evidence strongly suggests that generic drugs are not perceived to be perfect substitutes to the original brand-name drug, despite being chemically identical. After generic entry, the original brand-name firm typically charges a higher price than its generic version and still has positive market shares (e.g. Grabowski and Vernon (1992), Frank and Salkever (1997), and Scott Morton (2000)). These findings fit well with predictions of vertical differentiation models. Two recent papers applied to branded-generic competition are Cabrales (2003) and Königbauer (2005).

[4]One can think of the entry costs as a marketing cost associated with entering a new country-specific market. Alternatively, the entry costs can be thought of as (expected) R&D costs which must be recouped for the discovery of a new drug treatment to take place.

[5]Pavcnik (2002) provides strong evidence from Germany that the introduction of reference pricing has

The inclusion of on-patent drugs is perhaps the main source of controversy over reference price systems. It is argued that TRP per se effectively eliminates patent protection and will stifle innovation in drug therapy, while GRP, on the other hand, is considered to have a minimal effect on incentives for R&D, since it applies only to off-patent drugs (see e.g. Danzon (2001) and Lopez-Casasnovas and Puig-Junoy (2000)). The model confirms the first line of the argument, but not the second. TRP provides the lowest profits to the patent-holding firm, making market entry of the new drug treatment least likely.[6] However, a patent-holding firm can be negatively affected by reference pricing, even if on-patent drugs are exempted from this particular reimbursement system. Stronger price competition induced by GRP forces the patent-holding firm to lower the price of its drug in order to reduce the loss of market shares.

Another important concern about TRP is that this system forces a large number of patients to opt for a less suitable drug simply to avoid the extra co-payment. The broader the therapeutic cluster, the more severe is the trade-off between surcharges and increased health risks to patients.[7] GRP, on the other hand, is said to conserve third party funds without exposing patients to significant risks, because it applies to substitution only among generically equivalent drugs that have demonstrated bioequivalence to the original brand-name drug. For given prices, this is, of course, trivially true. However, the intention of the reference price systems is to induce price responses from the pharmaceutical firms. Taking this into account, it can be shown that, in fact, GRP distorts drug choices most and exposes patients to higher health risks. Since the on-patent drug is exempted from reference pricing under GRP, the patent-holding firm faces a less price-elastic demand than its competitors and can

induced pharmaceutical prices to drop, the effect being stronger for branded drugs facing generic competition. Aronsson et al. (2001) provide similar evidence from Sweden.

[6]This result has empirical support from Danzon and Ketcham (2004) who analyse the effect of reference pricing on the availability of drugs in Germany, the Netherlands, and New Zealand.

[7]Lopez-Casasnovas and Puig-Junoy (2000, p. 111) formulate this problem as follows:

> "First, if there is no interchangeability at the level of the individual patient [...] then the co-payment may become not avoidable and the reference price system may discriminate against some patients. Second, selection of a drug under a reference price category may result in a lower level of effectiveness and potentially harmful side effects for the patient because the drug is chosen simply with a view to avoiding the copayment."

The same argument is presented by Danzon (2001).

thus charge a considerably higher price. This induces a larger fraction of patients to choose the drugs that are included in the reference cluster, which are less suitable, but have a lower co-payment.

In terms of policy implications, the results suggest no clear-cut conclusions about the optimal choice of the reimbursement system. However, distinctions among the following general cases can be made. If the risk of no market entry for new drugs is low, i.e. market entry costs are low and the expected return in the (country-specific) market is high, then TRP is clearly socially favourable. However, if this is not the case, then either NRP or GRP might be necessary to stimulate market entry. The choice between NRP and GRP implies a trade-off, since the former yields higher drug expenditures but lower health risks to patients. A social planner's evaluation of this particular trade-off is determined by the importance of the drug expenditures in the planner's objective function. Hence, GRP might be the favoured reimbursement system in countries where the pharmaceutical industry is insignificant or non-existent, while NRP might be preferred otherwise.

The theoretical literature on reference pricing is, as mentioned above, very limited with only a couple of notable exceptions. Zweifel and Crivelli (1996) analyse the pricing responses to the introduction of a reference price system using a Bertrand duopoly model. They frame their analysis in the context of the introduction of the TRP system in Germany in 1989. Danzon and Liu (1998) use a monopolistic competition model with kinked demand and imperfect physician agency to predict price responses to reference pricing. The modelling approaches are distinctly different from the model in this chapter. With the combination of horizontal and vertical differentiation, GRP and TRP can be analysed and compared closely. Moreover, the analysis of market entry and health risks to patients is possible which are lacking in the above mentioned studies.[8]

This chapter contributes also to the more general literature on horizontal and vertical product differentiation. Most papers within this field allow firms to invest in quality, but assume consumers to differ only in terms of the horizontal space (taste).[9] The present

[8]These important aspects of reference price systems are also absent in Merino-Castelló (2003) who studies the price effects of generic reference pricing in a vertical differentiation model.

[9]Several papers have added quality competition to a standard Hotelling-framework, see e.g. Ma and Burgess (1993) for the case of fixed locations under both price competition and price regulation, Economides (1989) for the case of endogenous locations and price competition, and Brekke et al. (2006) for the case of

model explicitly combines the horizontal differentiation framework of Hotelling (1929) with the vertical differentiation framework introduced by Gabszewicz and Thisse (1979, 1980) and Shaked and Sutton (1982, 1983). While these two approaches typically are applied separately, the pharmaceutical market – with both inter-brand (branded vs. branded) and intra-brand (branded vs. generic) competition – serves as a natural example for combining these frameworks.

The chapter is structured as follows: In section 3.2, the model is presented. In section 3.3, the equilibrium prices are derived and characterised for all three regimes. Section 3.4 analyses the market entry decision of the firm with the new drug treatment. Section 3.5 discusses the welfare properties of the three different regimes and presents some policy implications. Finally, in section 3.6, the chapter is concluded.

3.2 The Model

Consider a particular therapeutic market for prescription drugs with the following characteristics. There are two patient types, indexed by $j = H, L$, differing with respect to their gross valuation of drug treatment due to, e.g., different degrees of illness. A fraction λ of the patients are H-types with a gross valuation v. The remaining patients, the L-types, have a gross valuation γv, where $\gamma \in (0, 1)$. Both patient types are uniformly distributed on the line segment $S = [0, 1]$, with a total mass of 1, where the location of an arbitrary patient, $x \in S$, is associated with the patient's susceptibility towards specific drug characteristics. A 'mismatch cost' parameter t measures the utility loss per unit of distance between a patient's ideal treatment, given by his location on S, and the drug which is actually consumed. These mismatch costs can reflect various side-effects or contraindications that reduce the gross valuation of the drug treatment.

There are potentially three pharmaceutical single-product firms, indexed by $i = 0, 1, G$, operating in the market. Firms 0 and 1 offer original brand-name drugs at prices p_0 and p_1, respectively. These drugs, which differ with respect to chemical compounds, are located

endogenous locations and price regulation. However, none of these papers allow consumers to differ with respect to their willingness-to-pay for quality, which means that the vertical differentiation framework is not explicitly dealt with.

at either end of the unit interval S, reflecting their horizontally differentiated treatment effects. It is assumed that drug 1 is a new treatment version, which is still under patent protection, and that will be introduced in this particular market, if the variable profits are sufficient to cover the entry costs. Drug 0, on the other hand, has already lost its patent protection and faces generic competition from the third pharmaceutical firm G which offers a generic drug version at a price p_G. In terms of horizontal differentiation, the generic drug is (naturally) also positioned at 0. However, in the eyes of the patients, 0 and G are vertically differentiated. This is captured by assuming that the patients' gross valuation of the generic drug is deflated by a factor $\theta \in (0,1)$. Thus, the perceived quality difference between the two versions of drug treatment 0 is given by $(1-\theta)$. This vertical differentiation might be due to differences in advertising intensity that creates perceived quality differences, or simply due to the brand-name drug being perceived to be safer, because of its longer life in the market.

Each patient needs one unit of either drug version. A patient of type j who is located at x and consumes a unit of drug i obtains the following utility:

$$U_j(x,i) = \begin{cases} u_j - t|x-i| - c_i & \text{if} \quad i = 0,1 \\ \theta u_j - tx - c_i & \text{if} \quad i = G \end{cases} \tag{3.1}$$

where

$$u_j = \begin{cases} v & \text{if} \quad j = H \\ \gamma v & \text{if} \quad j = L \end{cases} \tag{3.2}$$

c_i is the patient's co-payment for drug i.

Patients are (partially) insured and face a co-payment rate $\alpha \in (0,1)$. In the absence of a reference price system, the co-payment is simply given by $c_i = \alpha p_i$. In the presence of a reference price system, the co-payment is based on a reference price $\overline{p}$, and the patients must pay the full price difference, if they choose a drug in the reference group which is priced in excess of the reference price. Hence, if drug i is included in a reference price system, the co-payment is given by

$$c_i = \begin{cases} \alpha p_i & \text{if} \quad p_i \leq \overline{p} \\ \alpha \overline{p} + (p_i - \overline{p}) & \text{if} \quad p_i > \overline{p} \end{cases} \tag{3.3}$$

The model is a three-stage game with the following sequence of events:

1. A benevolent regulator decides on the socially optimal drug reimbursement policy to implement. He chooses among the following policies:

(i) no reference pricing (NRP),

(ii) therapeutic reference pricing (TRP), or

(iii) generic reference pricing (GRP).

2. Firm 1 decides whether to enter the market and thus to offer a new treatment, given that treatment 0 already exists and is offered in the form of both an original version (drug 0) and a generic substitute (drug G).

3. All pharmaceutical firms in the market play a simultaneous pricing game.

As usual, the game is solved by backward induction.

3.3 Drug Pricing

In this section, the optimal pricing strategies of the pharmaceutical firms are derived for each of the three possible reimbursement regimes. In equilibrium, all firms are active and compete in terms of prices. This requires some restrictions on the parameters. More specifically, the mismatch cost parameter t must be bounded from both below and above, i.e. $t \in \left(\underline{t}, \overline{t}\right)$, where the lower and upper bounds are functions of the other parameters. In the Appendix (6.2.1-6.2.3), it is shown that, when $t \in \left(\underline{t}, \overline{t}\right)$, there exists a vertically separating equilibrium, where the brand-name drug 0 is priced 'high' and consumed by the H-types only, while the generic substitute G is priced 'low' and consumed by the L-types only.[10] The horizontally differentiated brand-name drug 1 is consumed by both types in equilibrium. This is the only possible type of equilibrium, where the generic drug can survive in the market, since all patients prefer drug 0 over drug G, if $c_0 = c_G$, implying that either all or no patients of type j prefer 0 over G, if $c_0 \neq c_G$.

Proposition 1 *In equilibrium, the brand-name drug 0 is priced higher than its generic substitute, if* $t \in \left(\underline{t}, \overline{t}\right)$. *The H-patients consume the brand-name version, whereas the L-patients consume the generic version.*

[10] To be more precise, an equilibrium exists when $t \in \left(\underline{t}, \overline{t}^k\right)$, $k = NRP, TRP, GRP$. In other words, there is a common lower bound on t in all three regimes, whereas the upper bound generally differs between the regimes.

It is worth noting that, in this context, it makes intuitive sense to focus on intermediate values of the mismatch cost parameter t. On the one hand, a very low t is not compatible with patent protection, since a new drug must be sufficiently differentiated to obtain a patent. On the other hand, a very high t is not compatible with the notion of a 'therapeutic market'. In particular, the idea of therapeutic reference pricing requires that the drugs included in a reference group are not too differentiated.

Demand and profits

The drug demand is derived for each firm under the assumption of vertical market segmentation. This requires the identification of two indifferent patients – one for each of the two patient types.

The H-types choose between the two brand-name drugs, and the location of the indifferent H-type patient, denoted $\widetilde{x}_H$, is given by the solution to

$$U_H\left(\widetilde{x}_H, 0\right) = U_H\left(\widetilde{x}_H, 1\right)$$

yielding

$$\widetilde{x}_H = \frac{1}{2} + \frac{c_1 - c_0}{2t} \tag{3.4}$$

The L-types, on the other hand, choose between the generic drug G and the horizontally differentiated brand-name drug 1. The location of the indifferent L-type patient, denoted $\widetilde{x}_L$, is given by the solution to

$$U_L\left(\widetilde{x}_L, G\right) = U_L\left(\widetilde{x}_L, 1\right)$$

yielding

$$\widetilde{x}_L = \frac{1}{2} + \frac{c_1 - c_G - \gamma v\left(1 - \theta\right)}{2t} \tag{3.5}$$

Under the additional assumption of full market coverage, so that all patients obtain non-negative utility from the consumption of their most preferred drug, the demand facing firm i is given by

$$D_i = \begin{cases} \lambda \widetilde{x}_H & \text{if} \quad i = 0 \\ \lambda\left(1 - \widetilde{x}_H\right) + \left(1 - \lambda\right)\left(1 - \widetilde{x}_L\right) & \text{if} \quad i = 1 \\ \left(1 - \lambda\right)\widetilde{x}_L & \text{if} \quad i = G \end{cases} \tag{3.6}$$

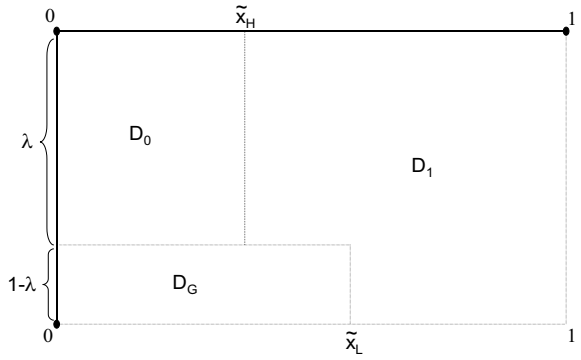

Figure 3.1: Illustration of the demand system

Figure 3.1 illustrates the demand system. Firm 0, providing the old breakthrough drug, serves only the high valuation (high severity) patients, represented by a fraction λ, whereas the generic firm serves the low valuation (low severity) patients, given by the fraction $1-\lambda$. The new on-patent drug producer serves both segments.

Finally, assuming zero production costs, the (variable) profits for firm i are simply given by[11]

$$\pi_i = p_i D_i \tag{3.7}$$

3.3.1 No Reference Pricing

In the absence of any reference price system, a patient's co-payment for drug consumption is simply given by

$$c_i^{NRP} = \alpha p_i^{NRP} \tag{3.8}$$

Explicit expressions for the profit functions under the NRP system are easily found by using (3.8) in (3.4)-(3.7). In equilibrium, the two brand-name producers choose the prices p_0^{NRP} and p_1^{NRP} that maximise π_0 and π_1, respectively, as defined by (3.7). The optimal strategy for the generic producer, on the other hand, is to choose a price p_G^{NRP} that is just low enough to make it unprofitable for firm 0 to deviate from p_0^{NRP} by setting a 'low' price that also

[11] At this stage, market entry costs (R&D costs and/or marketing costs) are sunk and thus play no role for the analysis.

captures the L-types. The equilibrium drug prices are given by[12]

$$p_0^{NRP} = \frac{3t}{\alpha}\Delta_0 \tag{3.9}$$

$$p_1^{NRP} = \frac{t}{\alpha}\Delta_1 \tag{3.10}$$

$$p_G^{NRP} = \frac{1}{\alpha}\left[3t\Delta_G - \gamma v\left(1-\theta\right)\right] \tag{3.11}$$

where

$$\Delta_0 := \frac{3-(1-\lambda)\sqrt{1-\lambda}}{8+\lambda\left(\lambda^2+3(1-\lambda)\right)} > 0 \tag{3.12}$$

$$\Delta_1 := \frac{10-\lambda\left(\lambda^2+3(1-\lambda)\right)-6\left(1-\lambda\right)\sqrt{1-\lambda}}{8+\lambda\left(\lambda^2+3(1-\lambda)\right)} > 0 \tag{3.13}$$

$$\Delta_G := \frac{4-\lambda(2-\lambda)-(4-\lambda)\sqrt{1-\lambda}}{8+\lambda\left(\lambda^2+3(1-\lambda)\right)} > 0 \tag{3.14}$$

All prices are increasing in t and decreasing in α. Higher mismatch costs reduce the substitutability and thus the degree of competition between the brand-name drugs, leading to higher prices. A higher co-payment rate, on the other hand, increases the price-elasticity of drug demand, leading to lower prices in equilibrium. It is also straightforward to show that $\partial\Delta_i/\partial\lambda > 0$, implying $\partial p_i/\partial\lambda > 0$, for all $i = 0, 1, G$. A higher fraction of H-types implies an increase in the overall willingness to pay with a corresponding price increase for the original drugs. This price increase also enables the generic producer to charge a higher price in equilibrium.[13] Note also that a reduction of the perceived quality difference between the two versions of treatment 0, i.e. an increase in θ, leads, as expected, to a higher price for the generic drug version.

A higher gross valuation of drug treatment for the L-types, i.e. an increase in γ, leads, counterintuitively, to a lower generic price in equilibrium. The reason is that a higher gross valuation for the L-types, which implies a higher willingness-to-pay for drugs, makes it more profitable for firm 0 to lower its price in order to capture the L-segment of the market. Consequently, the generic firm must reduce its price in order to prevent this price-undercutting strategy from the brand-name firm. If the difference in gross valuations between the two patient types becomes sufficiently small, i.e. if γ becomes sufficiently close to 1, it

[12] A full derivation of the equilibrium is given in the Appendix (6.2.1).

[13] From (3.11) and (3.14), note that λ must be sufficiently high to secure a non-negative generic drug price and thus equilibrium existence. See the Appendix for the exact conditions.

is not possible for the generic firm with a (perceived) lower-quality product to prevent the brand-name firm from serving both patient types in equilibrium. In this case, the generic drug is driven out of the market.

From (3.9)-(3.11), the following ranking of the equilibrium drug prices can be established:

$$p_0^{NRP} > p_1^{NRP} > p_G^{NRP} \tag{3.15}$$

These price differences are reflected in the allocation of the equilibrium market shares[14]:

$$\widetilde{x}_H^{NRP} = \frac{3\left[3-(1-\lambda)\sqrt{1-\lambda}\right]}{2\left[8+\lambda\left(\lambda^2+3(1-\lambda)\right)\right]} \in \left(\frac{3}{8},\frac{1}{2}\right) \tag{3.16}$$

$$\widetilde{x}_L^{NRP} = \frac{3\left[2+(2-\lambda)\lambda+(2+\lambda)\sqrt{1-\lambda}\right]}{2\left[8+\lambda\left(\lambda^2+3(1-\lambda)\right)\right]} \in \left(\frac{1}{2},0.77\right) \tag{3.17}$$

Proposition 2 *Under NRP, the brand-name drug with a generic substitute always charges the highest price in equilibrium. Both patient groups are distorted. H-type patients consume more of the new, patent-protected brand-name drug, while L-type patients consume more of the generic drug.*

It might seem counterintuitive that the price level is highest for the brand-name drug with a generic substitute, since, normally, prices would be expected to be lower for products that face stronger competition. The reason for this result is that, due to generic competition, the optimal strategy of firm 0 is to concentrate exclusively on serving the H-type patients and leave the L-types to the generic competitor. Since firm 0 competes only for H-patients with a less price-elastic demand, while firm 1 competes for both patient types, firm 0 sets a higher price than firm 1 in equilibrium. This theoretical result tallies well with several empirical findings of price increases for brand-name drugs after the entry of generic substitutes in the market.[15]

[14]These equilibrium market shares can then be compared to the allocation $\widetilde{x}_j = \frac{1}{2}$ that minimises the patients' mismatch costs.

[15]The empirical study by Grabowski and Vernon (1992) shows that generic entry was followed by price increases by the branded producer, a result later confirmed by Frank and Salkever (1997). This finding was called the 'generic competition paradox' by Scherer (1993).

Inserting the equilibrium prices into (3.7), the equilibrium profits can be derived:

$$\pi_0^{NRP} = \frac{3t\lambda\Delta_0}{2\alpha}\left(1+\Delta_1-3\Delta_0\right) \tag{3.18}$$

$$\pi_1^{NRP} = \frac{t\Delta_1}{2\alpha}\left(1+3\left(\Delta_G\left(1-\lambda\right)+\lambda\Delta_0\right)-\Delta_1\right) \tag{3.19}$$

$$\pi_G^{NRP} = \frac{1-\lambda}{2\alpha}\left(1+\Delta_1-3\Delta_G\right)\left(3t\Delta_G-\gamma v\left(1-\theta\right)\right) \tag{3.20}$$

3.3.2 Reference Pricing

Consider now the implementation of a reference price system. Some drugs are aggregated into a cluster and are subject to the same reference price $\overline{p}$. The introduction of a reference price system involves the following decision-making.

First, the regulator must decide which drugs to include in a cluster or reference group. In the present model, this choice boils down to whether or not the new brand-name drug should be included. Inclusion of the horizontally differentiated new drug implies therapeutic reference pricing (TRP). On the other hand, if the reference group consists only of the old brand-name drug and its generic substitute, the reimbursement system is characterised as generic reference pricing (GRP).

Second, the regulator must decide on the reference price level. In most countries, this level is set at, or close to, the lowest drug price in the cluster. In the present analysis, this practice is followed by assuming that the lowest price in the reference group, i.e. the generic price, is chosen as the reference price level: $\overline{p}=p_G$.

Therapeutic Reference Pricing

Under TRP, the reference group consists of all three drugs in the therapeutic market, i.e. also the horizontally differentiated drug 1. By the assumption of $\overline{p}=p_G$, the co-payments faced by the patients under TRP are given by

$$c_i^{TRP}=\begin{cases} p_i^{TRP}-\left(1-\alpha\right)p_G^{TRP} & \text{if} \quad i=0,1 \\ \alpha p_G^{TRP} & \text{if} \quad i=G \end{cases} \tag{3.21}$$

The co-payments differ as compared to NRP, since the patients that are prescribed one of the original drugs are now also fully liable for the price difference with respect to the reference price.

As before, explicit expressions for the profit functions under the TRP system are found by using (3.21) in (3.4)-(3.7), and the derivation of the equilibrium is similar to that under the NRP system (see Appendix 6.2.2). The equilibrium prices under TRP are given by

$$p_i^{TRP} = \alpha p_i^{NRP}, \quad i = 0, 1, G \tag{3.22}$$

Thus, compared with NRP, TRP implies that the prices are set, *as if* $\alpha = 1$. The reason is that, with TRP, the patients are fully liable for any price increase above the reference level. This also implies that the equilibrium prices are independent of the co-payment rate. Furthermore, since the equilibrium market shares are independent of α, both patient types are equally distorted under the two regimes.

Proposition 3 *In equilibrium, the relative price differences and market shares are equal under NRP and TRP.*

Compared with the NRP case, the (uniform) downward pressure on the drug prices under TRP is also reflected in lower equilibrium profits, now given by

$$\pi_i^{TRP} = \alpha \pi_i^{NRP}, \quad i = 0, 1, G \tag{3.23}$$

Generic Reference Pricing

Under GRP, only generic substitutes are grouped into the same cluster as the original, off-patent drugs. Horizontally differentiated, but therapeutically equivalent drug versions are not included. Hence, the co-payments faced by the consumers under GRP are given by

$$c_i^{GRP} = \begin{cases} p_i^{GRP} - (1-\alpha)\, p_G^{GRP} & \text{if} \quad i = 0 \\ \alpha p_i^{GRP} & \text{if} \quad i = 1, G \end{cases} \tag{3.24}$$

While only a fraction α of the drug price needs to be paid on the drugs G and 1, patients that are prescribed the brand-name drug 0 must additionally pay the full price difference between the original drug and the reference price.

The equilibrium prices, derived in the same way as previously (see Appendix 6.2.3), are

given by

$$p_0^{GRP} = \frac{\left(2+\alpha-\sqrt{1-\lambda}\,(2-\lambda-\alpha)\right)\Gamma}{\widetilde{\Delta}} \tag{3.25}$$

$$p_1^{GRP} = \frac{t\overline{\Delta}+(1-\alpha)(1-\theta)\gamma v\left(\widehat{\Delta}-2(2+\alpha)\right)-\sqrt{1-\lambda}\,(2\alpha-\lambda(\alpha+1))\Gamma}{\alpha\widetilde{\Delta}} \tag{3.26}$$

$$p_G^{GRP} = \frac{3t(\alpha\lambda-3\lambda+\lambda^2+4)-\gamma v(1-\theta)\widehat{\Delta}-(4-\lambda)\Gamma\sqrt{1-\lambda}}{\widetilde{\Delta}} \tag{3.27}$$

where

$$\widehat{\Delta} := 4\alpha+5\lambda-2\alpha\lambda-4\lambda^2+\lambda^3+\alpha\lambda^2+4>0 \tag{3.28}$$

$$\widetilde{\Delta} := 8\alpha+8\lambda-6\alpha\lambda-5\lambda^2+\lambda^3+2\alpha\lambda^2+\alpha^2\lambda>0 \tag{3.29}$$

$$\overline{\Delta} := 10\alpha+\lambda-6\alpha\lambda+2\lambda^2-\lambda^3+\alpha\lambda^2+2\alpha^2\lambda>0 \tag{3.30}$$

$$\Gamma := 3t-\gamma v(1-\theta)(1-\alpha)>0 \tag{3.31}$$

Using the equilibrium prices derived above, the equilibrium market shares under GRP are characterised by the location of the indifferent patient in each patient-group:

$$\widetilde{x}_H^{GRP} = \frac{\Gamma\left[(2+\alpha)-(2-\lambda-\alpha)\sqrt{1-\lambda}\right]}{2t\widetilde{\Delta}} \tag{3.32}$$

$$\widetilde{x}_L^{GRP} = \frac{\Gamma\left[\alpha(2-\lambda)+\lambda(3-\lambda)+(2\alpha+\lambda)\sqrt{1-\lambda}\right]}{2t\widetilde{\Delta}} \tag{3.33}$$

Comparing with (3.16)-(3.17), it is also relatively straightforward to verify that

$$\widetilde{x}_j^{GRP}>\widetilde{x}_j^{TRP}=\widetilde{x}_j^{NRP},\quad j=H,L \tag{3.34}$$

implying that more patients choose one of the drugs included in the reference cluster under GRP, i.e. drug 0 and G.

In order to evaluate the ranking of the equilibrium prices under GRP, a rather weak assumption on the co-payment rate is needed, namely that $\alpha<\frac{2}{3}$. Then, the pricing equilibrium under GRP can be characterised as follows:[16]

$$p_1^{GRP}>p_0^{GRP}>p_G^{GRP} \tag{3.35}$$

[16]For $\alpha\geq\frac{2}{3}$, the ranking of p_0^{GRP} and p_1^{GRP} is ambiguous. It is possible to derive the exact condition for this price ranking to hold for $\alpha\geq\frac{2}{3}$, but the condition is rather cumbersome and also hard to interpret, so the analysis focuses on the plausible case of $\alpha<\frac{2}{3}$. The exact condition can, however, be provided by the authors upon request.

Proposition 4 *Assume that* $\alpha < \frac{2}{3}$. *Then, under GRP, the brand-name firm without a generic substitute always charges the highest price in equilibrium. Both patient groups are generally distorted. The L-types always consume more of the generic drug, while the H-types consume more of the new patent-protected brand-name drug, if* λ *and/or* t *are sufficiently low, and more of the old off-patent product otherwise.*

A proof is given in the Appendix (6.2.4).

The ranking of the equilibrium prices changes under a GRP system. The price is now higher for the brand-name drug without a generic substitute. The reason is simply that drug 1 is not included in the reference cluster. If a consumer chooses this drug, the co-payment is given by a share α on the *total* drug price. In contrast, if the off-patent drug 0 is chosen, which is included in the reference cluster, the full price difference between the generic substitute and the brand-name drug must be paid. Thus, by not having its product included in the reference group, firm 1 faces a less price-elastic demand and will consequently charge a higher price in equilibrium.

In contrast to the NRP or TRP systems, the equilibrium price differences do not automatically translate into equivalent differences in the equilibrium market shares. This is due to the asymmetry which is introduced by different co-payments for patients and which depends on whether or not the demanded drug is subject to reference pricing. Therefore, even if firm 1 sets the highest drug price, it may not be the most expensive alternative for the consumers and, consequently, this firm may have a higher market share in the H-segment. Proposition 4 shows that this is the case, if λ and/or t are sufficiently low. In this case, the price of the on-patent drug is kept relatively low in order to capture a larger share of the L-segment (which is more important, the lower the level of λ) and/or is due to fierce competition induced by a relatively low degree of horizontal differentiation.

On the other hand, the location of the indifferent L-type patient is always distorted towards drug 1, as before. In other words, due to the price difference between the generic and brand-name drugs, a larger share of L-patients consumes the generic drug G. Finally, it should be noted that even though the H-segment may be distorted 'both ways' under GRP, the L-segment is always more distorted towards drug 1. This can easily be verified from (3.32)-(3.33) by confirming that $\widetilde{x}_L^{GRP} > \widetilde{x}_H^{GRP}$.

Using the equilibrium prices reported in (3.25)-(3.27), the equilibrium profits under GRP can be derived. These profit expressions are rather tedious, and are therefore relegated to the Appendix (6.2.5).

3.4 The Market Entry Decision

When interpreting the market in question as *country-specific* therapeutic market, demarcated by national regulation, it can be assumed realistically that firm 1 will only enter this particular market, i.e. offer its newly developed product in this country, if the expected profits from sales in this market cover the market entry costs. When considering the costs and benefits of entry, the firm must take into account, how the reimbursement policy in a given country is likely to affect the profits from drug sales in this country.

There is a clear-cut ranking of the equilibrium profits for the potential entrant (firm 1) across the different reimbursement regimes:

Proposition 5 *The equilibrium profits of the patent-holding entrant are always highest under NRP and lowest under TRP.*

A proof is given in the Appendix (6.2.6).

The profit comparison between NRP and TRP is straightforward. Compared with the case of NRP, the TRP system puts a downward pressure on drug prices, while keeping equilibrium market shares intact. This implies that the profits are unambiguously lower in the TRP equilibrium. NRP also outperforms GRP, from the viewpoint of firm 1, since prices *and* market shares are higher in the former case. A comparison between GRP and TRP, on the other hand, shows that in the former case prices are higher, but market shares lower. Nevertheless, the equilibrium profits are always higher under GRP. The reason is that under GRP, firm 1 faces a drug demand with a lower price-elasticity, which enables the firm to charge a considerably higher price while suffering only a moderate loss of market shares. All else equal, it follows that the expected profits for a potential entrant are always lowest, when entering a market that is subject to TRP, and highest when entering a market with NRP.

This result is not surprising and tallies well with the popular concern about TRP with

respect to a potential erosion of patent rights, as discussed in the Introduction. However, it is worth noting that a patent-holding firm can be negatively affected by reference pricing, even if on-patent drugs are exempted from this particular reimbursement system. In the model, firm 1's profits are lower under GRP compared to NRP, even though drug 1 is not included in the reference cluster. The reason is that firm 1 offers a drug that is an imperfect substitute to the drugs directly affected by the GRP system. Stronger price competition between the firms 0 and G, induced by GRP, implies that firm 1 is also forced to lower the price of its on-patent drug in order to reduce the loss of market shares.

3.5 Welfare Analysis

In this section, it is discussed, how considerations for social welfare will influence the optimal choice of the reimbursement scheme for pharmaceuticals. It is assumed that a regulator has two main concerns:

(i) Minimisation of the aggregate mismatch costs C_k as a measure of the patients' total health risks from drug consumption.

(ii) Minimisation of total drug expenditures E_k.

The weighting of these two objectives, in case of conflict, might depend on the characteristics of the country in question. Two polar cases can be distinguished:

In countries with a significant pharmaceutical industry, it is reasonable to assume that the profits of pharmaceutical firms matter for the national regulator. In this case, the welfare costs of higher drug prices might be restricted to the efficiency costs of the increased third-party funding for drug expenditures.[17] Naturally, a regulator will put relatively more emphasis on minimising aggregate mismatch costs from drug consumption in this case (this will also be the case, if the perspective of global welfare is taken).[18] On the other hand,

[17]In a model with unit demand, the patients' copayments for drugs are an efficient transfer from consumers to producers with no efficiency costs associated.

[18]Note that it is implicitly assumed that market entry is always socially optimal, i.e. that market entry costs are low compared to the induced mismatch costs without the horizontally differentiated treatment

in countries with no pharmaceutical industry, it is reasonable to assume that drug expenditures are more important in terms of national welfare. Indeed, a stated desire behind the introduction of reference pricing in many countries is precisely to curb the total outlays on pharmaceutical consumption.

In the subsequent analysis, first the effect of the different reference price systems on the aggregate mismatch costs is examined, and then the effect on drug expenditures, given that there is market entry by firm 1. The optimal choice of the reimbursement system in the presence of the partly conflicting regulatory goals is then discussed in the next section, where the effect of the reimbursement choice on the market entry decision is taken into account.[19]

3.5.1 Mismatch Costs

The total mismatch costs under the reimbursement system k, denoted by C_k, are given by

$$C_k = \lambda \left[\int_0^{\widetilde{x}_H^k} st \; ds + \int_{\widetilde{x}_H^k}^1 (1-s)\, t \; ds \right] + (1-\lambda) \left[\int_0^{\widetilde{x}_L^k} st \; ds + \int_{\widetilde{x}_L^k}^1 (1-s)\, t \; ds \right] \tag{3.36}$$

Clearly, the total mismatch costs are minimised, if $\widetilde{x}_L^k = \widetilde{x}_H^k = \frac{1}{2}$. In other words, mismatch costs are minimised, if all patients located at $x \leq \frac{1}{2}$ are prescribed either drug 0 or G, while all patients located at $x > \frac{1}{2}$ are prescribed drug 1. However, due to price differences, total mismatch costs will never be minimised in equilibrium. It was previously shown that $\widetilde{x}_j^k \neq \frac{1}{2}$ for at least one patient type in all three reimbursement regimes. Furthermore, the equilibrium market shares are equal under NRP and TRP, implying that total mismatch costs must also be equal under these two regimes.

The explicit expression for the total mismatch costs in each of the three different regimes, which are quite tedious, are given in the Appendix (6.2.7). Based on these expressions, the following unambiguous ranking of the reimbursement systems with respect to the equilibrium

version. This assumption can be rationalised, because both treatment versions 0 and 1 must be sufficiently differentiated for drug 1 to receive patent protection ($t > \underline{t}$).

[19] In the discussion of the welfare and the policy implications, the implicit assumption is made that a regulator does not take the 'artificial' vertical differentiation between the branded and generic drugs into account, but rather attaches the same gross utility to objectively homogeneous products.

mismatch costs can be derived:[20]

Proposition 6 *NRP and TRP yield equal mismatch costs in equilibrium. These are always lower than under GRP.*

In order to explain this result, consider the distortive effects of GRP on each of the two patient types. Note that $\widetilde{x}_L^{GRP} > \widetilde{x}_L^{TRP} = \widetilde{x}_L^{NRP} > \frac{1}{2}$, due to the larger price difference between the generic drug and the horizontally (and vertically) differentiated drug 1 under GRP. This implies that GRP always increases the total mismatch costs in the L-segment. For the H-types, on the other hand, $\widetilde{x}_H^{TRP} = \widetilde{x}_H^{NRP} < \frac{1}{2}$ and $\widetilde{x}_H^{GRP} > \widetilde{x}_H^{TRP} = \widetilde{x}_H^{NRP}$. However, since $\widetilde{x}_H^{GRP}$ can be larger or smaller than $\frac{1}{2}$, it is possible that GRP reduces the aggregate mismatch costs for the H-types, if $\widetilde{x}_H^{GRP}$ is sufficiently close to the midpoint of the line segment S. Nevertheless, a possible reduction in the mismatch costs for the H-types will always be dominated by the increase in the mismatch costs for the L-types. The reason is twofold. First, the mismatch costs are only reduced for the H-types, if λ, the fraction of the H-types in the population, is sufficiently low (see Proposition 4). In this case, the contribution of the H-types to the *total* mismatch costs is also relatively low. Second, since the location of the indifferent L-type is further away from the midpoint of S in all regimes, the effect of a marginal relocation of the indifferent patient on the total mismatch costs is, all else equal, larger in the L-segment.

The result stated in Proposition 6 is perhaps somewhat surprising. It certainly runs contrary to the popular concern about the discriminatory effects of TRP, that this reimbursement system forces a larger number of patients to opt for a less suitable drug simply to avoid the extra co-payment and thereby increasing the mismatch costs. However, this is not the case in the model. True, TRP will increase the overall mismatch costs *for given prices*, if we use the NRP case as a benchmark. But this argument ignores the fact that the pharmaceutical firms will adjust their pricing policies according to the drug reimbursement system. Here, we have seen that TRP will lead to a proportionally equal reduction in all drug prices, leaving the patients' drug choices unaffected in equilibrium compared to NRP.

[20]The proof, though conceptually straightforward, involves some extremely tedious algebra and is thus not reported. However, just to give a brief sketch, it is possible to show that $C_{GRP} - C_{NRP} = \frac{\varphi_1}{\varphi_2}$, where $\varphi_2 > 0$ and φ_1 is a convex quadratic function of t which crosses zero from below at $t = \underline{t}$. Thus, $\varphi_1 > 0$ for $t > \underline{t}$. It follows that $C_{GRP} > C_{TRP} = C_{NRP}$ for $t > \underline{t}$.

GRP, on the other hand, will lead to more distorted drug choices due to larger equilibrium price differences within the therapeutic market. Since the on-patent drug is exempted from reference pricing under GRP, firm 1 faces a less price-elastic demand than its competitors and can thus charge a considerably higher price in equilibrium. This, in turn, induces more patients to choose the drugs that are included in the reference cluster. This leads to higher overall mismatch costs.

3.5.2 Drug Expenditures

In order to evaluate, how the different reimbursement systems affect drug expenditures, first the equilibrium price levels for the same drugs across different regimes must be compared. Using the equilibrium prices reported for the different cases above, it is relatively straightforward to verify that

$$p_i^{NRP} > p_i^{GRP} > p_i^{TRP}, \quad i = 0, 1, G \tag{3.37}$$

for all $t > \underline{t}$. This result reflects and confirms the main rationale behind reference pricing. By introducing a reference price system, price competition is generally increased, since the price-elasticity of drug demand increases for prices above the reference price level. This effect is the stronger, the more drugs are included in the reference cluster, implying that drug prices are lower under TRP than under GRP. Since prices are strategic complements, the introduction of a reference price system of either kind puts a downward pressure on the prices of *all* drugs in the market. Compared with the NRP case, the introduction of GRP has a direct negative effect on the price level of drug 0, which in turn leads to a reduction also in the price of drug 1, even though this drug is not included in the reference cluster under GRP. Furthermore, by going from GRP to TRP, firm 1 gets a direct incentive to cut its drug price, which then indirectly leads to a further price reduction also for drug 0. Finally, lower prices for brand-name drugs imply that the generic producer must also lower its price in order to stay in the market.

This price ranking reflects in an equivalent ranking of the drug expenditures:

$$E_{NRP} > E_{GRP} > E_{TRP} \tag{3.38}$$

for all $t > \underline{t}$. The expenditure comparison between GRP and TRP is straightforward. We know from Proposition 5 that firm 1's profits are higher under GRP than under TRP, because

prices are higher. This result holds despite the fact that the market shares $(1 - \tilde{x}_H)$ and $(1 - \tilde{x}_L)$ are both lower under GRP than under TRP. The expenditures for both drug 0 and drug G are then clearly lower under TRP, because both their market shares $\tilde{x}_H$ respectively $\tilde{x}_L$ and their prices are lower under TRP.
The expenditure comparison between NRP and GRP is less straightforward. Clearly, the expenditure for drug 1 is higher under NRP than under GRP. Again, this is due to the expenditures reflecting the profit ranking for firm 1. Despite the fact that firm 0's market share $\tilde{x}_H$ is smaller under NRP than under GRP, its profit is higher under NRP, because the price is higher under NRP and this positive price effect dominates the negative demand effect. The intuition is that firm 0 is less restricted in its optimisation problem under NRP, and its 'unrestricted' profits must thus be at least as high as its 'more restricted' profits. G's profit is also higher under NRP, because the positive price effect dominates the negative demand effect. By reducing the price level to the level under GRP, G could increase its market share $\tilde{x}_L$ to an even higher level than under GRP and reach thus a higher profit. However, since this is not optimal, G's profit under NRP must be higher.

Proposition 7 *The prices and equivalently the drug expenditures in the therapeutic market are highest under NRP and lowest under TRP.*

3.5.3 Policy Implications

If a regulator seeks to minimise overall mismatch costs, the above analysis suggests that GRP should never be implemented. The mismatch costs are minimised by choosing either NRP or TRP.

However, there are other considerations that might be taken into account. First, the price level of pharmaceutical drugs will play a role, if the regulator is concerned about curbing the total outlays on pharmaceuticals. As previously discussed, the relative weighting of mismatch costs and expenditures in the welfare function is likely to depend on the relative importance of the pharmaceutical industry in the country in question. The more important the pharmaceutical industry is, the less concerned a regulator should be about pharmaceutical prices. In any case, as long as the regulator places any weight on pharmaceutical prices at all, the above analysis clearly suggests that a TRP system should be implemented, since this reimbursement scheme minimises both mismatch costs and prices.

However, this conclusion is only valid, if there is indeed an additional, horizontally differentiated drug version that can be included in the therapeutic cluster. Since equilibrium profits are lowest under TRP, this reimbursement system makes market entry least likely for a given level of market entry costs. If the possibility of no market entry is taken into account, then the welfare considerations are no longer clearly in favour of TRP. First, no entry will lead to *maximal* mismatch costs, because only one treatment version (drug 0 and its generic substitute) is offered in the market. Second, the absence of competition from a horizontally differentiated drug will lead to increased drug prices and thus higher drug expenditures under both NRP and GRP. In this scenario, the regulator must take into account, how the choice of the reimbursement system is likely to affect the probability of market entry for new drugs.

No clear-cut conclusions can be made about the optimal choice of reimbursement systems. However, based on the above analysis, the following classification of scenaria can be made:

- TRP, which minimises both mismatch costs and drug prices, is clearly the socially favourable reimbursement system, if the risk of no market entry for new drugs is low. However, if this is not the case, then either NRP or GRP might be necessary to stimulate market entry.
- NRP minimises mismatch costs, but maximises drug prices. It might be optimal for countries where drug prices do not play an important role in the objective function of the regulator due to a dominant pharmaceutical industry.
- GRP, on the other hand, might be the favoured reimbursement system in countries where the pharmaceutical industry is insignificant or non-existent, since it leads to lower drug prices than NRP.

The probability of market entry clearly depends on the market entry costs which can be interpreted as bureaucracy costs, licensing costs, and so on, and are therefore sunk. But market entry furthermore depends on country-specific demand-side effects, since they define the expected return for firm 1. Hence, in countries where the return is expected to be high, TRP is more likely to be encountered than in countries where this is not the case.

This classification of scenaria suggests an optimal choice of the reimbursement system under different national circumstances concerning the risk of no market entry and the signif-

icance of the pharmaceutical industry. Clearly, there are various additional variables which determine the choice in reality and which were not considered in the model. Therefore, it is difficult to compare precisely the suggested choice with the implemented regime in practice. Danzon and Ketcham (2004) analyse the TRP-system in New Zealand, Germany, and the Netherlands, and some interesting similarities with the suggested results in the present model can be found:

Since New Zealand is not a very important market for the pharmaceutical industry, it faces a serious threat not to be provided with pharmaceutical products, if the regulation is too strict and the expected profits thus too low. The TRP-regime in New Zealand is very strict and, indeed, the number of provided drugs is significantly lower as compared to other industrialised countries.

Germany, on the contrary, represents a lucrative market with price-independent and rather low copayments for pharmaceuticals such that even a strict reference price system is alleviated. Hence, the most efficient regime for Germany seems to be TRP which was recently re-implemented after Germany changed from TRP to GRP in 1996.

The Netherlands should clearly consider the threat not to be provided with drugs due to TRP. But, empirically, not significantly fewer drug versions are found than, e.g., in Germany. The reason might be that they have a very lenient reference price level which counteracts the negative effect of TRP on the expected profits.[21]

Anecdotal evidence suggests that the countries which implemented GRP fit rather well with the predictions from the model. The pharmaceutical market in Sweden, Italy, Spain, and Danmark is supposedly not significant enough to avoid the risk of no market entry. Hence, they refrain from strict TRP. At least in Italy and Spain, the pharmaceutical industry is unimportant such that the model strongly suggests the implementation of GRP. Danmark's and Sweden's pharmaceutical industry is, however, more important. Other reasons seem to play an important role why low pharmaceutical prices are weighted more than the patients' increased health risks due to the mismatch costs.

[21] This reference price level was fixed at the generic price level in the present model.

3.6 Conclusion

The effects of different reference price systems for pharmaceuticals were analysed focusing on a specific therapeutic market with potentially three pharmaceutical firms. Two of the firms offer horizontally differentiated brand-name drugs. One of these drugs is off-patent and faces competition from a generic version offered by a third firm. The other drug is on-patent and will be introduced in the market, if the profits are sufficient to cover the entry costs.

Within this framework, generic reference pricing (GRP) and therapeutic reference pricing (TRP) could be compared, as well as the benchmark case of no reference pricing (NRP). TRP triggers competition most and results both in lower equilibrium prices for every drug in the therapeutic market and in the lowest health risks for the patients. In contrast, GRP results in higher prices and distorts drug choices most. Hence, GRP leads to a higher level of patient health risks, measured in terms of aggregate mismatch costs, than the other two reimbursement systems. TRP is thus preferable from the perspective of both the purchaser (payer) and the patients.

Notably, the beneficial role of TRP crucially relies on the assumption that the new on-patent drug enters the market. If the market entry costs are sufficiently high and/or the expected country-specific returns sufficiently low, TRP may in fact result in a worse outcome than both GRP and NRP, as described above. This leads to another interesting topic that was not addressed in this chapter:
It is often argued that TRP may induce pharmaceutical firms to invest more in drastic innovations, which are not subject to reference pricing, rather than me-too innovations, which very likely will be included in a reference group. The trade-off with respect to therapeutically similar me-too innovations is the following: On the one hand, they increase competition and lower pharmaceutical prices as well as reduce the patients' mismatch costs by offering a different variant of treatment for the same illness. On the other hand, me-too innovations might crowed out drastic innovations, if they reduce the total budget available for R&D. This is, however, only partly true, because different drug versions are often innovated in so-called R&D-races, meaning that me-too innovations are already in the 'pipeline' of innovations when the first drastic innovation enters the market. It is therefore questionable, whether they really crowd out drastic innovations, or whether they simply lost the R&D-race.

Chapter 4

Parallel Imports and Price Regulation

4.1 Introduction

In order to limit health expenses, countries have an incentive to regulate prices. However, parts of the pharmaceutical spending is needed to cover the enormous R&D costs that arise, before a drug even can be placed on the market.[1] Danzon (1997b) calculates that R&D accounts for roughly 30 percent of total costs. Grabowski and Vernon (1990) conclude that new drugs, that finally enter the market, typically earn at most modest profits, and 70 percent earn insufficient excess returns to cover the average cost of R&D (OECD Health Data (2003)). It is often argued that governments have an incentive to free-ride on those sunk innovation costs by setting drug prices equal to marginal costs. This is individually rational, since it minimises health costs, but it reduces or even eliminates the pharmaceutical firms' incentives to invest in R&D and leads to a suboptimal supply of new products. This result hinges on the fact that a closed economy is considered, where parallel imports are not allowed.

The rules concerning parallel imports vary across countries. The World Trade Organisation's Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) provi-

[1] For a survey on research and development costs of pharmaceutical firms see DiMasi et al. (1991, 2003).

sions permits individual countries to choose their own policies on exhaustion of intellectual property rights. The European Union, for example, applies the principle of regional exhaustion, which means that parallel imports from European countries are allowed. This has significant implications, if one considers that the European Union hosts both low-priced countries, such as France, Greece, and Spain, and high-priced countries, e.g. Germany, the United Kingdom, and the Netherlands. Arbitrage opportunities will result in a uniform price level within the European Union.[2] It is estimated that the share of parallel imports will increase from 8 percent of the pharmaceutical drug market in 2001 to 10 percent in 2006 (Atkinson (2001)).

In this chapter, it is shown that national price regulations will grant mark-ups on marginal costs, if parallel imports are allowed. Because price differences result in arbitrage opportunities, governments need to explicitly induce the pharmaceutical firms to export the products to their countries. For this, the price ceilings have to be sufficiently high.

There have been several studies arguing in general terms for or against parallel imports.[3] However, only a few theoretical models are actually applied to study the effects of parallel imports on pharmaceutical prices in detail, and the models reach different welfare implications.

Bordoy and Jelovac (2005) analyse the welfare effects of parallel imports using a two-countries model where the countries differ in their levels of copayments and in the patients' valuation of the drug. They conclude that parallel imports increase welfare, if countries differ in their patients' valuation, but that welfare decreases, if countries differ in their health system.

Danzon (1997b)[4] explores the conflict between marginal cost pricing and efficient investment in R&D and incorporates innovation into the welfare analysis. She argues that with R&D costs being globally joint and sunk, Ramsey pricing is second-best optimal (if some form of non-linear pricing is ruled out). But with parallel imports, manufacturers tend to set uniform prices. Formerly low-priced countries are the losers, since prices increase for them. In the long-run, even high-priced countries lose, because fewer drugs are invented overall. As a solution, she suggests exempting on-patent drugs from parallel imports or allowing

[2]These arbitrage opportunities have even been improved in 1996, when the European Medicines Evaluation Agency (EMEA) harmonised the regulatory requirements for drug approval, packaging, and labelling, which has reduced the importers' transaction costs.

[3]See e.g. Bale (1998), Chard and Mellor (1989), Maskus (2000), or Berndt (2002).

[4]See also Danzon (1997a, 1998), Danzon and Towse (2003), and Danzon et al. (2003) for further references.

confidential contracting with individual governments.

Similarly, Darba and Rovira (1998) argue for the European pharmaceutical market that different drug prices are welfare improving. They can be achieved by discounts or reference pricing, where in a system of reference pricing various drugs with similar treatment effects are grouped in a therapeutic class, for which one single price ceiling is set by the regulator. Malueg and Schwartz (1994) find that uniform pricing yields a lower global welfare than third-degree discrimination, if demand dispersion across markets is sufficiently large and vice versa. Mixed systems, in which blocks of countries with similar income levels permit parallel trade, yield greater benefits than either uniform pricing in all markets or discrimination. They argue that the European Union should put its member states into sub-groups, where parallel imports are only allowed within each of them.

Most closely related to this chapter is Pecorino (2002). He examines a partial equilibrium model of trade in which a monopolist can sell in his home and in a foreign country. The firm can set the monopoly price in the home country, whereas the foreign price is determined by a Nash Bargaining Game between the firm and the foreign government. Pecorino compares welfare in the home country, if reimports are allowed, and if they are not allowed. If reimports are allowed, the foreign price will be established in the home country as well. Therefore, the firm has an incentive to bargain even harder and it will only export to the foreign market, if overall profits are at least as high as the monopoly profit at home. He concludes that reimports always increase welfare in the home country with linear demand or with a general demand function where the foreign government has the whole bargaining power.

Following Pecorino (2002), the present model also assumes that the pharmaceutical firm can explicitly make an export decision.[5] However, it does not examine the firm's position, but rather takes the benevolent regulators' side in a global perspective and analyses their incentives to grant a mark-up on marginal costs. It can be shown that, if parallel imports are not allowed, importing countries have an incentive to use marginal cost pricing, whereas allowing for parallel imports results in a price ceiling between marginal costs and the home country's price ceiling. The firm earns positive profits. In the end, there are effectively no parallel imports, because by deterring them, the pharmaceutical firm can make higher profits. Furthermore, the model shows in an extension that, in this static setting, individu-

[5]Danzon et al. (2003) empirically showed that parallel imports and reference pricing indeed induced pharmaceutical firms to delay the introduction of new drugs, if they feared the spill-over of low prices.

ally regulating countries produce higher pharmaceutical drug prices than a single benevolent regulator.
With the number of foreign parallel-importing countries sufficiently high, all foreign countries set their price ceilings only slightly above marginal costs, thus taking advantage of the 'quantity effect' (all foreign countries together need to fulfil the participation constraint). The mark-up on marginal costs is determined by the home country, which sets the highest possible ceiling, i.e. the monopoly price. In the end, the pharmaceutical firm makes monopoly profits based on the home country's demand.

This chapter contributes to the ongoing debate about the advantages and disadvantages of allowing parallel imports. The model ignores dynamic effects on R&D and investigates instead static effects. The dynamic argument for prices above marginal costs is limited by a significant drawback, because it can basically explain any pricing in equilibrium. By restricting the analysis to static effects, the result of a price ceiling above marginal costs gains value, since it is much harder to achieve.

The chapter is structured as follows: In section 4.2, a simple two-countries model is presented to illustrate the basic effects. After a short presentation of the standard argument without parallel imports, the effect of parallel imports are dealt with. In section 4.3, several extensions are provided. First, it is shown that price regulation in a static setting is optimally allocated to the countries individually rather than to a single benevolent regulator who could not exploit the strategic pricing effect. Then, the analysis is generalised to n countries. In section 4.4, some policy implications are derived. Finally, section 4.5 concludes.

4.2 A Two-Countries Analysis

4.2.1 The Standard Argument

Standard theory, see e.g. Danzon (1997a), argues that countries tend to free-ride on the huge sunk innovation costs of pharmaceutical firms. By setting the drug ceiling as low as possible, they can minimise their health expenses. This seems to be rational, but this regulation has severe impacts on the pharmaceutical firm's incentive to invest in further pharmaceutical innovations. Ex ante, there are huge investments necessary to finally bring a new drug on

the market, and these investments are already sunk at the time when the national regulators decide on the price ceiling. In order to recover these R&D costs, a sufficiently high mark-up on the price must be granted.

In a static perspective, a classical free-rider problem arises, because the innovation costs are sunk. As long as the price covers at least the marginal costs of production, the firm provides the drug. Hence, all individually regulating countries have an incentive to implement marginal cost pricing. They hope that the *foreign* price regulations allow for a sufficiently high profit such that the sunk innovation costs are covered and further innovations can be expected despite their own strict price regulation.[6] In the long run, this is a problem, because the development of new drugs increases social welfare by allowing to treat illnesses more effectively or by providing new treatments. But the pharmaceutical firms anticipate the problem and expect losses, since the sunk innovation costs cannot be recovered with marginal cost pricing. In the extreme, there will be no investments in R&D.

This general argument can be shown in the framework of a two-countries model without arbitrage possibilities, where governments regulate prices sequentially. A pharmaceutical firm produces a drug q in country H. It faces large sunk innovation costs I and marginal costs c. The drug can be sold in the firm's home country (H) and a foreign country (F). Demand is $q_i(p_i)$, with $q_i'(p_i) = \frac{dq_i(p_i)}{dp_i} < 0$ and $i = H, F$. Demand depends on the price which eventually prevails in the market.

The structure of the model is as follows: In the first stage, the firm decides whether to innovate. The decision rule is to innovate, if the profits resulting from innovation are at least as high as the expected innovation costs. If the profits are lower, the firm does not invest. In the second stage, the H government sets a price ceiling which determines the maximal price allowed in H. The optimal price ceiling maximises welfare in H. Welfare consists of the consumer surplus

$$CS_H(p_H) = \int_{p_H}^{q^{-1}(0)} q_H(p)dp \tag{4.1}$$

[6]Precisely, the firm does not compare expected profits, but rather expected returns with the innovation costs. Since the innovation costs are sunk at the time of the profit respectively return realisation, this chapter talks about the profit π.

and the firm's profits[7]

$$\pi_i(p_i) = (p_i - c)q_i(p_i), \qquad i = H, F \tag{4.2}$$

Hence, H maximises the welfare function defined in (4.3) by choosing an appropriate price ceiling p_H.

$$\max CS_H(p_H) + \pi_H(p_H) + \pi_F(p_F) \tag{4.3}$$

The game is solved by backward induction.

In the third stage, F sets its optimal price ceiling. F's welfare consists only of the consumer surplus, because the pharmaceutical firm's profits accrue solely to the home country. F maximises its welfare function subject to the firm's participation constraint. The firm provides the country F only with the drug, if it earns non-negative profits.

$$\max CS_F(p_F) \;\text{ s.t. }\; \pi_F(p_F) \geq 0 \tag{4.4}$$

It can easily be seen that F indeed sets its price ceiling equal to marginal costs, i.e. as low as possible, and therefore free-rides on innovation costs. The lowest price ceiling is the price level p_F at which the participation constraint is binding:

$$\begin{aligned} &\tfrac{dCS_F(p_F)}{dp_F} = -q_F(p_F) \leq 0, \qquad \pi_F \geq 0, \qquad \pi_F \tfrac{dCS_F}{dp_F} = 0 \\ &\Rightarrow \quad p_F^* = c \end{aligned}$$

In the second stage, H anticipates F's behaviour. Since the home country's price ceiling has no impact on F's price regulation, H maximises welfare by considering only the consumer rent and the firm's profit in H. Welfare is thus maximised with marginal cost pricing.

$$\begin{aligned} &\tfrac{d[CS_H(p_H)+\pi_H(p_H)+\pi_F(p_F)]}{dp_H} = -q_H(p_H) + q_H(p_H) + (p_H - c)\tfrac{dq_H(p_H)}{dp_H} = 0 \\ &\Rightarrow\; p_H^* = c \end{aligned}$$

In stage 1, the pharmaceutical firm anticipates that the prices will be regulated too low, and that the innovation costs will not be recovered. Therefore, it will not innovate.

The only way to solve the free-rider problem is for the national regulators to commit ex ante credibly to a sufficiently high price mark-up. One possibility is, e.g., to build up

[7]In the trade literature, it is generally assumed that the welfare function includes the firm's profit both in the domestic and the foreign country. See e.g. Spencer and Brander (1983) or Krugman (1984). They explicitly use the idea that national governments wish to help domestic firms expanding market shares in profitable areas.

reputation. This can be done in a dynamic framework which is not further analysed in this chapter. The present model analyses the possibility to discipline national regulators in their price-setting behaviour by allowing for parallel imports.

4.2.2 Allowing for Parallel Imports

The structure of the model changes only slightly compared to the situation without parallel imports. In the first stage, the pharmaceutical firm decides whether to invest in innovations. In the second stage, the H government sets a price ceiling by maximising welfare, which consists of consumer surplus and the firm's profits in H and F. In the third stage, the F government sets its price ceiling, maximising welfare which consists only of the consumer surplus in F. But now, F needs to consider a different participation constraint of the firm: Since parallel imports are allowed, arbitrage opportunities resulting from price differences between H and F would induce parallel importers to reimport the drug to the high-priced country H.[8] The firm anticipates the potential parallel imports and sets the highest possible price level in the home country that just deters parallel imports. This is the foreign price level plus the transportation costs that a potential parallel importer would have to incur: $p_F + t$. The transportation costs are those costs that parallel-importing firms additionally incur for repackaging, relabelling, and so on. This price-setting strategy ensures the pharmaceutical firm higher profits than setting the maximal price ceiling p_H, because, by deterring parallel imports, it can serve the home market with the higher price $p_F + t$ rather than with p_F only. Hence, the participation constraint requires the overall profit from F and H at the lower price levels p_F respectively $p_F + t$ to be at least as high as the profit in H at the price p_H without exports:[9]

$$\text{(PC)} \qquad \pi_H(p_F + t) + \pi_F(p_F) \geq \pi_H(p_H) \tag{4.5}$$

In Arfwedson (2004), empirical evidence can be found that some pharmaceutical firms indeed think about restricting exports as a way to eliminate the possibility of parallel imports. GlaxoSmithKline stopped the exports to Canada in January 2003 in order to hinder the

[8] It will be shown later in the analysis that, with parallel imports being allowed, H sets a higher price ceiling than F.

[9] It is assumed that the transportation costs faced by the parallel importers are sufficiently low. If this were not the case, the threat of parallel imports would not be credible.

cheap reimports back into the USA. And already in 1991, Bayer tried to reduce the reimports from France and Spain to the United Kingdom by restricting its exports to France and Spain.

Stage 3: Price Regulation in F

F chooses its optimal price ceiling by maximising the consumer surplus in F subject to the new participation constraint of the firm.[10] The parallel importers' profits need not be included in the foreign welfare function. If there were parallel imports, Bertrand competition between them would result in zero profits. Therefore, they would not affect welfare.

$$\max CS_F(p_F) \text{ s.t. (PC) } \pi_H(p_F + t) + \pi_F(p_F) \geq \pi_H(p_H) \tag{4.6}$$

F has an incentive to set the price ceiling p_F as low as possible, because this maximises consumer surplus. Therefore, it chooses a price ceiling which satisfies the participation constraint with equality:

$$\begin{aligned} \text{(PC)} \qquad \pi_H(p_F + t) + \pi_F(p_F) &= \pi_H(p_H) \\ (p_F - c) \cdot [q_H(p_F + t) + q_F(p_F)] + t \cdot q_H(p_F + t) &= (p_H - c) \cdot q_H(p_H) \end{aligned}$$

The resulting reaction function $p_F(p_H)$ depends on the price level p_H that H sets in stage 2:

$$p_F^{PC}(p_H) = c + (p_H - c) \cdot \frac{q_H(p_H)}{q_H(p_F + t) + q_F(p_F)} - t \cdot \frac{q_H(p_F + t)}{q_H(p_F + t) + q_F(p_F)} \tag{4.7}$$

As long as the transportation costs t are sufficiently low, parallel imports improve thus the firm's bargaining position, and the foreign price ceiling lies between marginal costs and the home country's price ceiling: $p_F^{PC} \in (c, p_H)$.

Lemma 1 *The foreign country sets its price ceiling as low as possible. With parallel imports, the resulting foreign price ceiling is above marginal costs.*

Stage 2: Price Regulation in H

The pricing decision of H has a direct impact on the pricing decision of F, because it influences the participation constraint (PC). H anticipates this and maximises:

$$\begin{aligned} \max \quad & CS_H(p_F + t) + \pi_H(p_F + t) + \pi_F(p_F) = \\ & CS_H(p_F + t) + (p_F - c) \cdot [q_H(p_F + t) + q_F(p_F)] + t \cdot q_H(p_F + t) \end{aligned} \tag{4.8}$$

[10] A different participation constraint could also allow parallel imports from H to F, implying a lower threat to country F, if the drug were not supplied deliberately. The result does not differ qualitatively. Therefore, this case will not be investigated any further.

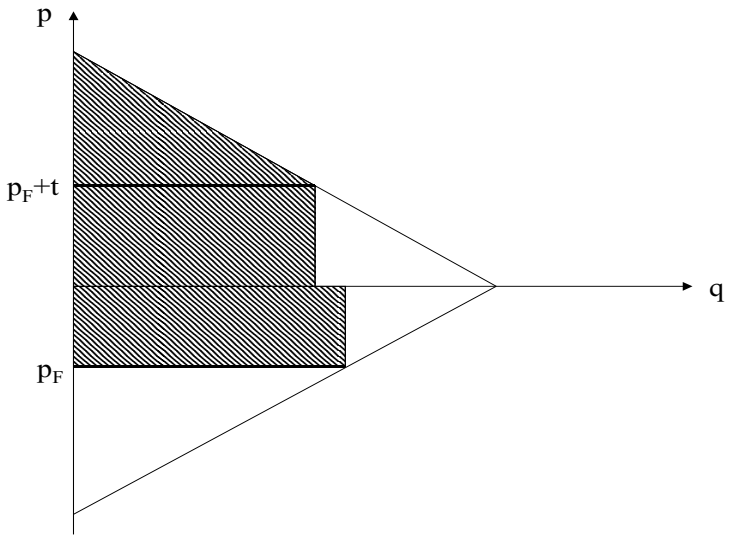

Figure 4.1: H's Maximisation Problem

Note that H takes the ex post emerging price level p_F respectively $p_F + t$ into account, with which, in the end, the consumers in the home country face a higher price than the consumers in the foreign country. The pharmaceutical firm deters all parallel imports by setting the price level equal to $p_F + t$ rather than applying the maximally allowed price ceiling p_H.

If H was able to set p_F directly, it would choose

$$p_F^* = c - \frac{q_F(p_F)}{q'_H(p_F + t) + q'_F(p_F)} - t \cdot \frac{q'_H(p_F + t)}{q'_H(p_F + t) + q'_F(p_F)} \tag{4.9}$$

with $q'_H(p_F + t) := \frac{dq_H(p_F+t)}{dp_F}$. The optimal price p_F^* trades off two opposing effects: If p_F increases, then welfare in H increases on the one side, because the pharmaceutical firm's profit rises. However, there is also an opposing effect. Welfare in H decreases on the other side, because the consumer rent falls. Starting from $p_F = c$ and thus from the social welfare's maximum, this negative effect is dominated by the former positive effect on profits. Therefore, the optimal foreign price ceiling from the viewpoint of H is above marginal costs.

H knows that F chooses its price to fulfil the firm's participation constraint with equality. It can therefore set p_H^* such that F exactly chooses the welfare-maximising p_F^* (see Figure

4.1):

$$\begin{aligned} p_F^* &= p_F^{PC} \\ p_H^* &= c + t \cdot \frac{q_H(p_F+t)}{q_H(p_H)} + \frac{q_F(p_F)+q_H(p_F+t)}{q_H(p_H)} \cdot \left[-\frac{q_F(p_F)}{q'_F(p_F)+q'_H(p_F+t)} - t \cdot \frac{q'_H(p_F+t)}{q'_F(p_F)+q'_H(p_F+t)} \right] \end{aligned} \tag{4.10}$$

Again, the optimal price p_H^* trades off several effects. The expression within brackets represents the same effects as the ones that p_F^* trades off: If the price ceiling in H is above marginal costs, welfare in H rises on the one side, because the profits increase, welfare falls on the other side, because the consumer rent decreases. An additional positive effect ($t \cdot \frac{q_H(p_F+t)}{q_H(p_H)}$) incorporates that an increase in p_H has only an indirect effect on the consumer rent, because p_F follows the price increase less than proportionally. Therefore, a price mark-up is less severe concerning the welfare in H.

However, there is a limit on the optimal foreign price ceiling p_F^*, above which H cannot influence F any more. This limit is reached with the monopoly price p_H^M. Above p_H^M, a further increase in p_H has no more impact on F's regulation, because F knows that the firm would never fully exhaust the allowed price range.

$$\begin{aligned} \pi_H(p_H) &= (p_H - c) \cdot q_H(p_H) \\ \Rightarrow \quad p_H^M &= c - \frac{q_H(p_H^M)}{q'_H(p_H^M)} \end{aligned} \tag{4.11}$$

If this limit is reached, then the price ceiling in H enters the participation constraint in stage 3 with $p_H = p_H^M$, and F exactly sets the price that fulfils this participation constraint:

$$\begin{aligned} p_F^{PC} &= c + (p_H^M - c) \cdot \frac{q_H(p_H^M)}{q_H(p_F+t)+q_F(p_F)} - t \cdot \frac{q_H(p_F+t)}{q_H(p_F+t)+q_F(p_F)} \\ p_F^{PC} &= c - \frac{[q_H(p_H^M)]^2}{q'_H(p_H^M) \cdot [q_H(p_F+t)+q_F(p_F)]} - t \cdot \frac{q_H(p_F+t)}{q_H(p_F+t)+q_F(p_F)} \end{aligned} \tag{4.12}$$

Lemma 2 *As long as the home country sets a price ceiling below the monopoly price, the foreign country's price ceiling depends positively on the home country's price ceiling. Thereafter, the home country cannot further influence the foreign price.*

When welfare consisting of consumer and producer rent is maximised, the lowest possible price level is in general optimal. Therefore, it might be surprising that the welfare-maximising price ceiling in H grants a mark-up on marginal costs. The reason is that the home country's

welfare consists of the foreign profit in addition. H's price ceiling disciplines the foreign country and induces it to grant a mark-up on marginal costs, too. H basically trades off a loss in the consumer rent at home and an increase in profits in both countries. Starting at marginal cost pricing and thus in the social welfare's maximum, an increase in the price reduces the consumer rent. But the increase in profits dominate this negative effect. It is optimal to increase the price ceiling p_H until the ex post induced price level p_F just equalises the gain in profits and the loss in consumer rents. Hence, to allow for parallel imports is an instrument for the home country to credibly grant the pharmaceutical firm a mark-up on marginal costs. This price will never prevail in the market.

Stage 1: The Pharmaceutical Firm's Innovation Decision

Anticipating the governments' price setting behaviour in the stages 2 and 3, the pharmaceutical firm decides in the first stage whether to invest in R&D or not.[11] It simply compares overall expected profits with expected innovation costs I. Again, there are two possible solutions.

If $p_H^* > p_H^M$, then there is a corner solution.

$$\begin{aligned} \pi(p_H^* > p_H^M) &= \pi_H(p_F^{PC} + t) + \pi_F(p_F^{PC}) \\ &= (p_F^{PC} - c) \cdot [q_H(p_F^{PC} + t) + q_F(p_F^{PC})] + t \cdot q_H(p_F^{PC} + t) \\ &= -\frac{[q_H(p_H^M)]^2}{q'_H(p_H^M)} \end{aligned} \tag{4.13}$$

The monopoly profit in country H is guaranteed as overall profit.

If $p_H^* \leq p_H^M$, then there is an inner solution and the welfare maximising p_F^* can be implemented.

$$\begin{aligned} \pi(p_H^* < p_H^M) &= \pi_H(p_F^* + t) + \pi_F(p_F^*) \\ &= (p_F^* - c) \cdot [q_H(p_F^* + t) + q_F(p_F^*)] + t \cdot q_H(p_F^* + t) \\ &= -\frac{q_F(p_F^*) \cdot [q_H(p_F^* + t) + q_F(p_F^*)]}{q'_H(p_F^* + t) + q'_F(p_F^*)} \end{aligned} \tag{4.14}$$

In both cases, the firm will invest in innovation, if $\pi \geq I$.

[11] Note that this is modelled as a yes/no-decision in this model. A variant would be to model a quality decision: If profits are expected to be high, large investments could guarantee drastic innovations, whereas small expected profits would only lead to so-called non-drastic 'me-too'-products.

Lemma 3 *Since the price ceiling is set above marginal costs, the firm makes positive profits.*

Note that, if the expected R&D costs are higher than the profits ($I > \pi$), it is better for both governments to allow for even higher prices in order to induce innovation. Ex ante, they could announce their willingness to increase the firm's profits even further, hoping for the firm to invest more in R&D. However, without the possibility to write binding contracts, this would be cheap talk. As soon as the firm has innovated, the innovation costs are sunk and there is no reason for the governments to increase prices any further than to the maximum of welfare.

Combining Lemma 1 – 3, Proposition 1 can be derived.

Proposition 1 *The pharmaceutical firm's participation constraint guarantees price ceilings above marginal costs, if parallel imports are allowed. The firm makes positive profits. The threat of parallel imports can be used to discipline the individually price-regulating countries, although, in equilibrium, no parallel imports will be observed.*

4.3 Extensions

4.3.1 A Single Regulator

It can be shown that, within this framework, a single benevolent regulator is worse than delegating the regulation to every individual country. A benevolent regulator would maximise the combination of the home country's and the foreign country's welfare by choosing the optimal price caps, anticipating that the lower price level of the two (plus possibly the transportation costs) will prevail in both markets due to the threat of potential parallel imports.
If he sets different price ceilings in both countries, then the optimisation problem is[12]

$$\max \; CS_F(p_F) + CS_H(p_F + t) + (p_F - c)q_F(p_F) + (p_F + t - c)q_H(p_F + t) \tag{4.15}$$

[12] Assume without loss of generality, that $p_F < p_H$.

with the corresponding first-order condition

$$(p_F - c)\underbrace{\frac{dq_F}{dp_F}}_{<0} + (p_F + t - c)\underbrace{\frac{dq_H}{dp_F}}_{<0} = 0 \tag{4.16}$$

In the case of equal price ceilings (p) in both countries, the regulator maximises

$$\max \; CS_F(p) + CS_H(p) + (p - c)\left[q_F(p) + q_H(p)\right] \tag{4.17}$$

with the corresponding first-order condition

$$(p - c)\left[\frac{dq_F}{dp} + \frac{dq_H}{dp}\right] = 0 \tag{4.18}$$

In both cases, it is optimal to implement marginal cost pricing. Since there is a welfare loss due to the positive transportation costs in the case of different price ceilings, welfare is maximised for equal price ceilings in both countries.

Proposition 2 *In a static setting, individually regulating countries implement higher pharmaceutical drug prices than a single benevolent regulator.*

The intuition is that a single benevolent regulator does not consider the strategic pricing effect. Whereas, in the analysis above, the regulator of the home country could credibly commit to a positive mark-up on marginal costs, because he considered word-wide profits, but the consumer rent only at home, this possibility is eliminated in the case of one single regulator who maximises world-wide profits *and* word-wide consumer rents.

Danzon (1997b) suggests that a single benevolent regulator might take the R&D costs into account and apply the Ramsey pricing rule. However, this is not credible in a static analysis like the present one. The regulator is assumed to be myopic and to maximise every single period's welfare. As shown above, welfare is maximised by setting the price equal to marginal costs. In this framework, it is therefore not appropriate to assume that a single benevolent regulator would apply the Ramsey pricing rule.
The Ramsey pricing rule applies, if there are quasi-fixed costs to be covered. Quasi-fixed costs are costs which cease as soon as the firm shuts down. Therefore, a regulator facing quasi-fixed costs might be interested in covering them in order to prevent a shut-down. The R&D costs of the pharmaceutical firm, however, are not quasi-fixed, but rather sunk and do not affect the shut-down or drug provision decision. As soon as the firm has invested, it will provide the product for any price covering marginal costs. Hence, a single regulator would not use the Ramsey rule of pricing in this framework, but rather marginal cost pricing.

4.3.2 An N-Countries Analysis

In what follows, n countries will be considered in order to generalise the analysis. It will be assumed that first the home country (1) sets a price ceiling to a newly developed drug, and then all other $(n-1)$ countries set their price ceilings simultaneously.[13] Assume first that there are only 3 countries. In stage 3, the countries 2 and 3 simultaneously decide what price ceiling to set. Since they maximise national consumer surplus, they want to set the lowest possible price which satisfies the firm's participation constraint (PC) with equality.

Without loss of generality, assume that country 3 decides upon the optimal price ceiling, taking as given that country 2 does the same:

- If country 2 sets a price $p_2 > p_f$, with p_f as the lowest possible price level such that both countries will be supplied given that country 2 sets p_2, it is optimal for country 3 to set $p_3 = p_f(p_2)$. Only then, the loss of consumer surplus in country 3 is the lowest possible, and the price level $p_f(p_2)$ respectively $p_f(p_2) + t$ will prevail in all countries eventually.
- Assume now that country 2 sets a price $p_2 = p_f$. Now, any price $p_3 \geq p_f(p_2)$ guarantees that country 3 will be supplied with the drug at the lowest possible price level $p_f(p_2)$ respectively $p_f(p_2)+t$. However, if transportation costs are positive $(t > 0)$, $p_3 = p_f(p_2)$ will be strictly better.

Since both countries face the same optimisation problem, they both choose p_f as price ceiling. The lowest possible price ceiling is defined by the participation constraint:

$$(p_f - c) \cdot [q_1(p_f + t) + q_2(p_f) + q_3(p_f)] + t \cdot q_1(p_f + t) = (p_1 - c) \cdot q_1(p_1)$$

$$p_f = c + (p_1 - c) \cdot \frac{q_1(p_1)}{q_1(p_f + t) + q_2(p_f) + q_3(p_f)} - t \cdot \frac{q_1(p_f + t)}{q_1(p_f + t) + q_2(p_f) + q_3(p_f)} \tag{4.19}$$

With the analogous interpretation as in the two-countries case.

For n countries, again with country 1 being the firm's home country, the lowest possible

[13] Assuming a sequential regulating game between the countries does not change the result, if national price regulations can be adjusted accordingly. If this is not the case, then the pharmaceutical prices and profits are even higher.

price level is defined by:

$$p_f(n) = c + (p_1 - c) \cdot \frac{q_1(p_1)}{q_1(p_f + t) + \sum_{i=2}^{n} q_i(p_f)} - t \cdot \frac{q_1(p_f + t)}{q_1(p_f + t) + \sum_{i=2}^{n} q_i(p_f)} \tag{4.20}$$

The denominator is larger than the nominator and the difference increases with n. For $n \to \infty$, the fractions approximate zero and the resulting price within the parallel-importing area approaches marginal costs. This result is very intuitive, because with every additional country, overall demand increases ('quantity effect'). Concerning the binding participation constraint of the firm, the profit must be guaranteed that the firm could make, if it remained in the home country. The larger overall demand gets, the smaller the n price ceilings can be in order to guarantee this profit. Eventually, demand is so high that the price ceilings approach marginal costs. The pharmaceutical firm should be indifferent to the number of participants, because its profit is the same regardless how many countries join and how low the price gets in the end.[14]

The other countries' behaviour is anticipated by country 1, i.e. the firm's home country, in stage 2. Country 1 knows that, regardless of its own price ceiling p_1, the price $p_f(n \to \infty)$ approximates marginal costs, implying that the consumer surplus in country 1 experiences only a small loss. However, country 1 also anticipates that its price ceiling p_1 has an effect on the overall profit of the firm. By setting p_1, country 1 has an effect on the guaranteed profit level via the participation constraint. Since both consumer surplus and profits enter the welfare function, and the consumers' rent is, with $p_f(n \to \infty)$ approximating marginal costs, nearly as high as possible, country 1 maximises its welfare by allowing for the monopoly price and thus by guaranteeing the highest profit possible.

$$\begin{aligned} S_1 &= CS_1(p_f + t) + \pi_1(p_f + t) + \sum_{i=2}^{n} \pi_i(p_f) \\ &= \int_{p_f(p_1)+t}^{q^{-1}(0)} q_1(p)dp + (p_1 - c)q_1(p_1) \qquad (4.21) \\ \frac{\partial S_1}{\partial p_1} &= -q_1(p_f(p_1) + t) \underbrace{\frac{\partial p_f}{\partial p_1}}_{\to 0} + q_1(p_1) + (p_1 - c)q_1'(p_1) = 0 \\ \Rightarrow p_1^* &= c - \frac{q_1(p_1)}{q_1'(p_1)} = p_1^M \qquad (4.22) \end{aligned}$$

[14]It is assumed that, without participation in the parallel-importing area, the countries regulate their prices and use marginal cost pricing (see section 4.2.1). Hence, the pharmaceutical firm would make zero profits.

Stage 1 does not change with respect to the 2-countries-analysis. The firm anticipates that it will make monopoly profits and innovates, if the expected monopoly profits are sufficient to cover the R&D-costs.

Proposition 3 *In equilibrium, all foreign countries (2 to n) set the lowest possible price ceiling $p_f(n)$ and just guarantee the firm's participation. The home country allows for the highest (monopoly) price level and ensures thus the highest profit possible.*

4.4 Policy Implications

The European Union

The model strongly supports the European policy to explicitly allow for parallel imports. It shows that it is less severe to delegate price regulations to the national governments than often assumed (e.g. Danzon (1997b)), because individually regulating countries implement higher drug prices than a single benevolent regulator. This surprising result suggests that, in practice, even the individually regulating member states of the European Union might set sufficiently high drug prices and innovation incentives can be maintained.

It is of course clear that a single European-wide regulation might improve on this situation in a dynamic setting. In practice, infinitely many periods need to be considered and there are certainly means for a single benevolent regulator to commit to a price-setting behaviour which might improve on the above described situation. However, the analysis highlights that the current situation, where every member state decides on its own pricing strategy, does not lead to an automatic breakdown of the market.

The Optimal Extent of the Parallel-Importing Area

An interesting implication can be drawn regarding the optimal number of participants in the parallel-importing area. The more parallel-importing countries are included, the lower the price ceiling that must be set eventually without impairing the pharmaceutical firm's profit. If these countries want to be provided with drugs, then they must adjust their price ceilings. A problem might be, of course, that this adjusted price ceiling is too high such that only a small group of consumers in these countries can afford this drug or only a limited

amount can be provided by the national health authorities. However, the quantity effect must not be neglected, which ensures that the overall price ceiling decreases when the area is extended. This has effects both on the European Union, specifically with respect to its expansion, and on developing countries.

The model implies that the expansion of the European Union by several Eastern European countries in May 2004 does not constitute such a severe threat for the pharmaceutical industry as generally assumed. It was feared that including further countries into the parallel-importing area reduces the expected profits, because only the home-country's monopoly profit is guaranteed. It should, however, not be forgotten that, given price regulation is implemented in these countries, there is the free-rider problem implying zero profits in these countries without inclusion in the parallel-importing area.[15]

Developing countries often do not have sufficient financial means to pay the same price for drugs as industrialised countries do. Therefore, they cannot provide their patients with necessary drugs, although these drugs do exist and could be produced at nearly no cost. A possible strategy would therefore be to supply developing countries with drugs at a price without mark-up on production costs and to let the richer industrialised countries pay for the R&D-costs. Although this argument is generally very intuitive, pharmaceutical firms do not follow this strategy in reality. It is prevented by the firms' fear of creating arbitrage opportunities between the low-priced developing countries and the high-priced industrialised countries, even if parallel imports were illegal.
The model shows that the consequences for the developing countries, i.e. not to be supplied with drugs, need not be as drastic as suggested. The more countries are part of an area, where parallel imports are allowed, the lower will the final drug price be. This means that nations should not try to prevent parallel imports, but rather actively foster them with as many other (price-regulating) nations as possible. Including developing countries, which will cause a large quantity effect given their large population, will therefore lead to a lower acceptable drug price than excluding developing countries.
Overall, it cannot be said, of course, whether the resulting price level will be sufficiently low for the developing countries to afford the drugs. Again, this is a question which must be

[15]The European expansion as a natural experiment is a good opportunity to check the conclusions of the model. Due to the sharp institutional change, the pharmaceutical prices before and after as well as the extent of parallel imports can be compared as soon as enough data are available.

addressed empirically.

The Locational Choice of the Firm

In order to draw some policy implications with respect to the locational choice of the pharmaceutical firm, the model needs to be slightly enriched by a further stage. First, the firm decides on the optimal country in which to locate. Then, the game evolves as described above. After the locational choice, the firm decides whether to innovate or not, anticipating that the home country will set a price ceiling guaranteeing the monopoly profit (under the assumption that a sufficiently high number of nations are part of the reimporting area), and the finally resulting price level is chosen by the foreign countries such that it just guarantees the provision with the drug. The monopoly profit in the home country is:

$$\pi_1^M = (p_1^M - c)q_1(p_1^M) = -\frac{[q_1(p_1^M)]^2}{q_1'(p_1^M)} \tag{4.23}$$

In the first stage, before the firm actually innovates, it can decide in which country to locate. This country is the firm's 'home country' and sets the price ceiling which influences the other countries in their pricing decision. The firm chooses a nation in which the monopoly profit is highest. Clearly, this is the case where demand is high and/or where the price-elasticity of demand $\mid \eta \mid$ is low:

$$\pi_1^M = -\frac{[q_1(p_1^M)]^2}{q_1'(p_1^M)} = \frac{q_1 \cdot p}{-\frac{dq_1}{dp} \cdot \frac{p}{q_1}} = \frac{q_1 \cdot p}{\mid \eta \mid} \tag{4.24}$$

Demand is in general high where the income is sufficiently high and patients can afford to buy necessary drugs. Since developing countries lack sufficient financial means, the model predicts that pharmaceutical firms will not choose to locate in developing countries, but rather in industrialised countries where a higher income level guarantees a higher monopoly profit.[16] The choice of location between industrialised countries, e.g. within the European Union, is probably mostly guided by the second effect, the price-elasticity of demand. The price-elasticity of demand predicts by how much demand decreases, when the price increases marginally. Hence, the pharmaceutical firm chooses the country where a price change has the least effect on demand. But a price change does not have full impact on demand, anyway, if individuals have some sort of health insurance. The price-elasticity depends on the insurance's coinsurance rates. Pharmaceutical firms will therefore settle down in those

[16]Unfortunately, this aspect cannot be checked empirically, because no developing country is part of a reimporting area.

industrialised countries where copayment rates are low. This strategy guarantees the highest monopoly profit possible.

This means that research-intensive firms base their locational choice not only on supply-side regulations (such as price ceilings), but also on demand-side regulations (such as coinsurance rates or physicians' budgeting). Pharmaceutical firms should thus be observed in countries where both demand and supply is rather weakly regulated. Kanavos (1998, 2002) describes this tendency empirically. Within the European Union, the bulk of research is provided by the United Kingdom, Germany, and Danmark. These are also the countries where price regulations are not very restrictive and pharmaceutical firms can decide about their prices rather freely. With fixed contributions to drugs, the United Kingdom and Germany have rather price-inelastic and high levels of demand. The research activities in Greece, Portugal, Finland, Austria, and Belgium, in contrast, are negligible, since they basically concentrate on generics. Their price restrictions can be classified as rather high.[17]

Consumer Surplus

The home country's consumers must pay a slightly higher price due to the transportation costs and are therefore worse off than the consumers of all other countries. In order to reduce this disadvantage, the transportation costs should be kept as low as possible:

$$CS_1(p_f + t) = \int_{p_f+t}^{q^{-1}(0)} q(p)dp \quad \Rightarrow \quad \frac{dCS_1}{dt} = -q(p_f + t) < 0 \tag{4.25}$$

This was first done in 1996, when the European Medicines Evaluation Agency (EMEA) harmonised the regulatory requirements for drug approval, packaging, and labelling. This harmonisation was feared to introduce further losses to the pharmaceutical industry, but it could be shown within this framework that it does not affect profits, but rather increases consumer welfare in the pharmaceutical firms' home countries. With the number of participating countries sufficiently high, the pharmaceutical firm can anticipate a guaranteed monopoly profit based on the home country's demand independent of any transportation costs.

[17] Of course, there are also several countries like Sweden and France, where both rigid price regulations and a significant pharmaceutical industry are observed. However, these countries differentiate in their regulations between national and foreign firms.

4.5 Conclusion

It was shown that the free-riding problem, concerning the huge sunk innovation costs of pharmaceutical firms, need not be as severe as generally assumed. In fact, allowing for parallel imports, which themselves are said to discourage pharmaceutical innovations, alleviates this problem significantly. The regulating countries are disciplined by the possibility of parallel imports, because they must keep in mind that the lowest price will prevail in all countries and that the firm must therefore be induced to export. This is achieved by price ceilings above marginal costs. The exact mark-up is determined by the price ceiling in the firm's home country which is in general above marginal costs, assuming that the regulator maximises welfare consisting of both consumer surplus and the overall profit of the firm.

Although it was shown that the pharmaceutical profits are positive with parallel imports being allowed, even higher profits could certainly be imagined. If firms were allowed to set different prices for each country without parallel imports, they could exploit the differences in the demand structure of the various countries and increase their profits by price-discrimination. These profits would, of course, be much higher than the above derived monopoly profit which only depends on the demand in the home country. However, this is not the benchmark that must be considered in this case. In order to reduce the enormous expenses in the health market, most countries regulate pharmaceutical drug prices. This means that pharmaceutical firms cannot price-discriminate optimally, but need to adhere to national price ceilings. When the individual countries set these price ceilings, the free-riding problem arises which leads to inefficiently low ceilings, if parallel imports are forbidden. The relevant benchmark is therefore the situation in which the firm must adhere to inefficiently low price ceilings set by the different nations without parallel imports, rather than the situation in which the firm can price-discriminate optimally.

Chapter 5

Concluding Remarks

Many countries struggle with their health care system and especially with the financing of the health care expenses. It is virtually impossible to find the 'best' design that satisfies all involved parties, because good health and the restoration of health in the case of illness is basically seen as some sort of human right which must be guaranteed. There are indeed treatments for numerous medical conditions available, however, the problem lies in the financing. On the one hand, the relevant treatment is often too expensive for the sick person, on the other hand, it is often questionable from an objective, unconcerned point of view, whether it is justifiable to shift the expenses to the solidarity respectively to the insurance pool. This is the major problem that most societies nowadays face and which must be solved by the health authorities.

Due to the sensitive nature of the problem, it is important to find a structured approach that guarantees as much objectivity as possible, and I strongly believe that health economics can contribute to the decision making process in this respect.

Of course, the health care system and with it the related problems are vast. Therefore, I chose *one* area, the pharmaceutical market, in order to contribute to the ongoing debate. In my opinion, the pharmaceutical market with its numerous aspects to consider is an especially interesting area to analyse. The main focus of this doctoral thesis was thereby to balance the two contradicting aims of keeping up the firms' innovation incentives on the one side, and reducing the financial burden on the other side, where this trade-off gains even more significance in the global context.

I am certainly aware that this doctoral thesis could only address some aspects without providing a 'patent remedy' for all health care related problems. However, I hope that it contributes to the progress in which eventually a solid fundament for the national health care systems will be established.

Chapter 6

Appendix

6.1 Advertising and Generic Market Entry

6.1.1 Proof of Proposition 1(i)

It must be shown under which conditions $p_2^{B*}(k) > p_2^{G*}(k)$ forms a Nash Equilibrium. The equilibrium candidate results in the following returns:

$$R_2^{B*}(k) = \bar{t}^2(1-\delta)(\theta-1)k\left[\frac{k+2\theta-1}{3k+4(\theta-1)}\right]^2 \tag{6.1}$$

$$R_2^{G*}(k) = \bar{t}^2(1-\delta)(\theta-1)(k+2\theta-1)\left[\frac{2-k}{3k+4\theta-4}\right]^2 \tag{6.2}$$

It forms a Nash-Equilibrium, if neither B nor G has an incentive to deviate from $p_2^{B*}(k)$ respectively $p_2^{G*}(k)$, given that the competitor does not deviate.

(i) B's incentive to deviate

Given that G sets $p_2^{G*}(k)$, the brand-name firm might want to reduce its price to $p_2^B = p_2^{G*}(k)$ in order to serve the not-detailed market, too.[1] The demand for the generic drug is then zero, whereas the demand for the brand-name drug increases to the deviation demand

$$\widehat{D}_2^B(k) = k\left(\bar{t} - \frac{p_2^{G*}(k)}{\theta}\right) + (1-k)\left(\bar{t} - p_2^{G*}(k)\right) \tag{6.3}$$

[1]It is assumed that the physician prescribes the brand-name version, if he is indifferent between the two drug versions.

The deviation return for B is then

$$\begin{aligned}\widehat{R}_2^B(k) &= (1-\delta)p_2^{G*}(k)\cdot\widehat{D}_2^B(k)\\ &= \bar{t}^2(1-\delta)(\theta-1)(2-k)\frac{2(\theta-1)(\theta-k)+k\theta^2(3-k)+k^2(2\theta-1)}{\theta(3k+4\theta-4)^2} \quad (6.4)\end{aligned}$$

B does not deviate from $p_2^{B*}(k)$, if

$$\begin{aligned}R_2^{B*}(k) &\geq \widehat{R}_2^B(k)\\ 0 &\leq k(k+2\theta-1)^2-\frac{1}{\theta}(2-k)\left[2(\theta-1)(\theta-k)+k\theta^2(3-k)+k^2(2\theta-1)\right] \quad (6.5)\end{aligned}$$

This condition only depends on k and θ. When k and/or θ are sufficiently small, then B can earn positive returns by deviating and charging a lower price.

(ii) G's incentive to deviate

The only potentially profitable deviation for G is to concentrate on the not-detailed market with a higher price. There are two possibilities: G calculates the optimal price level that it would set in the not-detailed market alone. If this price level is higher than $p_2^{B*}(k)$, then it sets $p_2^{B*}(k)-\epsilon$ ($\epsilon\to 0$). If it is lower than $p_2^{B*}(k)$, then he sets this optimal price level $\widehat{p}_2^G$.

The optimal price level in the not-detailed market alone is found by maximising

$$\begin{aligned}\widehat{R}_2^G(k) &= (1-\delta)(1-k)p_2^G\left(\bar{t}-p_2^G\right) &(6.6)\\ \Rightarrow\ \widehat{p}_2^G &= \frac{1}{2}\bar{t} &(6.7)\\ \widehat{p}_2^G \geq p_2^{B*}(k) \ \Leftrightarrow\ \theta &\leq \frac{1}{4}\left[5-k+\sqrt{k^2+10k+1}\right] &(6.8)\end{aligned}$$

Case 1: $\theta \leq \frac{1}{4}\left[5-k+\sqrt{k^2+10k+1}\right]$

If $\widehat{p}_2^G \geq p_2^{B*}(k)$, then G sets the highest possible price level $p_2^G = p_2^{B*}(k)-\epsilon$, ($\epsilon\to 0$). The deviation return for G is then

$$\begin{aligned}\widehat{R}_2^G(k) &= (1-\delta)p_2^{B*}(k)(1-k)\left(\bar{t}-p_2^{B*}(k)\right)\\ &= \bar{t}^2(1-\delta)(\theta-1)(1-k)\frac{(k+2\theta-1)(7\theta-k\theta+4k-2\theta^2-5)}{(3k+4\theta-4)^2} \quad (6.9)\end{aligned}$$

G does not deviate from $p_2^{G*}(k)$, if

$$\begin{aligned}R_2^{G*}(k) &\geq \widehat{R}_2^G(k)\\ 0 &\leq (k+\theta-1)(2-k)^2-(1-k)(k+2\theta-1)(7\theta-k\theta+4k-2\theta^2-5) \quad (6.10)\end{aligned}$$

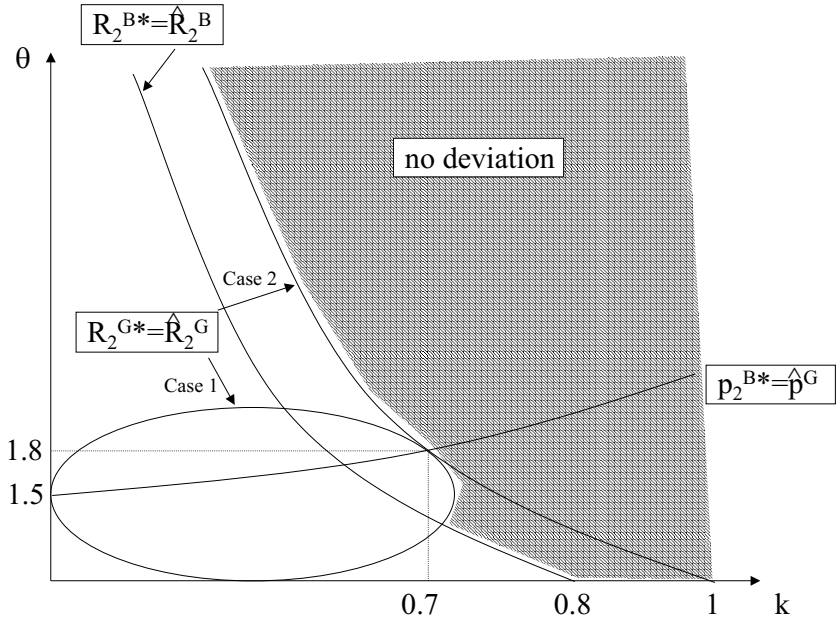

Figure 6.1: The Conditions for this Equilibrium.

Again, the condition only depends on k and θ.

Case 2: $\theta > \frac{1}{4}\left[5 - k + \sqrt{k^2 + 10k + 1}\right]$

If $\widehat{p}_2^G < p_2^{B*}(k)$, then G sets the optimal price level $\widehat{p}_2^G$. The deviation return for G is then

$$\widehat{R}_2^G(k) \;=\; (1-\delta)\widehat{p}_2^G(1-k)\left(\bar{t} - \widehat{p}_2^G\right) = \frac{1}{4}(1-\delta)(1-k)\bar{t}^2 \tag{6.11}$$

Although $\widehat{p}_2^G < p_2^{B*}(k)$, there is no demand for G from the detailed market:

$$\widehat{D}_2^G(k) \;=\; k\left[\frac{p_2^{B*}(k) - \widehat{p}_2^G}{\theta - 1} - \widehat{p}_2^G\right] = -\frac{1}{2}k\bar{t}\frac{2(\theta-1) + k(\theta+2)}{(\theta-1)(3k+4\theta-4)} < 0 \tag{6.12}$$

G does not deviate from $p_2^{G*}(k)$, if

$$\begin{aligned} R_2^{G*}(k) &\geq \widehat{R}_2^G(k) \\ \theta &\geq \frac{1}{2k}\left[-k^2 + 2 + \sqrt{k^4 - 5k^3 + 9k^2 - 8k + 4}\right] \end{aligned} \tag{6.13}$$

Figure 6.1 illustrates the conditions on k and θ that must be fulfilled in order for $p_2^{B*}(k) > p_2^{G*}(k)$ to be a Nash Equilibrium. A sufficient condition is (6.13).

Both k and θ must be sufficiently high for $p_2^{B*}(k) > p_2^{G*}(k)$ to be a Nash Equilibrium. Since k is an endogenous variable, whose optimal value is determined ex ante, it must be checked whether the profit-maximising k^* fulfils the derived conditions.[2]

6.1.2 Proof of Proposition 5(i)

It must be shown that $p_2^{B*} = \bar{p}$ and $p_2^{G*}(\bar{p}) = \frac{k\bar{p}+\bar{t}(1-k)(\theta-1)}{2(k+\theta-1)}$ form a Nash Equilibrium.[3] This is the case, if neither B nor G has an incentive to deviate, given that the competitor does not deviate.

(i) B's incentive to deviate

In the candidate equilibrium, the incumbent's second-period return is:

$$\begin{aligned} R_2^B(\bar{p}) &= (1-\delta)k\bar{p}\left[\bar{t} - \frac{\bar{p} - p_2^{G*}}{\theta - 1}\right] \\ &= \frac{1}{2}(1-\delta)k\bar{p}\frac{\bar{t}(\theta-1)(k+2\theta-1) - \bar{p}\,(k+2\theta-2)}{(k+\theta-1)\,(\theta-1)} \end{aligned} \tag{6.14}$$

The incumbent might have an incentive to deviate and to set $p_2^B = p_2^{G*}(\bar{p})$ in order to serve both markets completely. The deviation return is then

$$\begin{aligned} \widehat{R}_2^B &= (1-\delta)p_2^{G*}\left[k\left(\bar{t} - \frac{p_2^{G*}}{\theta}\right) + (1-k)\left(\bar{t} - p_2^{G*}\right)\right] \\ &= \frac{1}{4}(1-\delta)\left[\bar{t}(\theta-1)(1-k) + k\bar{p}\right]\frac{\bar{t}\left[k\theta(\theta(2-k)+2k-1) + k(1-k) + \theta(\theta-1)\right] - kp\left[k+\theta(1-k)\right]}{(k+\theta-1)^2\,\theta} \end{aligned} \tag{6.15}$$

B does not deviate, if $R_2^B(\bar{p}) \geq \widehat{R}_2^B$. This is the case for $\bar{p} \in [\bar{p}(1), \bar{p}(2)]$, i.e. the price cap is

[2] In the numerical illustrations, that are presented in section 2.3.3, all conditions are fulfilled and $p_2^{B*}(k) > p_2^{G*}(k)$ is a Nash Equilibrium.

[3] In order to facilitate the presentation, the dependence on k is not replicated.

bounded from below and above. The limits are defined by

$$\bar{p}(1) := \frac{\bar{t}(\theta-1)\left[k^3\left(\theta(\theta-1)+1\right)+k^2(\theta^2-1)+k\theta(\theta-1)(2\theta-1)-(k+\theta-1)\sqrt{\lambda}\right]}{k\left[5k\theta(\theta-1)+4\theta(\theta-1)^2+k^2(1+\theta^2)\right]} \tag{6.16}$$

$$\bar{p}(2) := \frac{\bar{t}(\theta-1)\left[k^3\left(\theta(\theta-1)+1\right)+k^2(\theta^2-1)+k\theta(\theta-1)(2\theta-1)+(k+\theta-1)\sqrt{\lambda}\right]}{k\left[5k\theta(\theta-1)+4\theta(\theta-1)^2+k^2(1+\theta^2)\right]} \tag{6.17}$$

$$\lambda := k\theta\left[k^3\theta+4k^2\left(1+2\theta(\theta-1)\right)+4k\left(\theta^2(\theta-2)+(\theta-1)\right)-4\theta(\theta-1)\right]$$

(ii) G's incentive to deviate

The generic second-period return in the equilibrium candidate is

$$\begin{aligned} R_2^G &= (1-\delta)p_2^{G*}\left[k\left(\frac{\bar{p}-p_2^{G*}}{\theta-1}-p_2^{G*}\right)+(1-k)\left(\bar{t}-p_2^{G*}\right)\right] \\ &= \frac{(1-\delta)\left[\bar{t}(1-k)(\theta-1)+\bar{p}k\right]^2}{4\left(k+\theta-1\right)\left(\theta-1\right)} \end{aligned} \tag{6.18}$$

If G deviates, it optimally sets the price level $p_2^G = \frac{1}{2}\bar{t}$ which is the optimal price level for the not-detailed market alone.

Case 1: $\bar{p} > \frac{1}{2}\bar{t}$

If the price cap is not binding, then G sets the optimal price level $\widehat{p}_2^G = \frac{1}{2}\bar{t}$. The deviation return is then

$$\widehat{R}_2^G = (1-\delta)(1-k)\widehat{p}_2^G(\bar{t}-\widehat{p}_2^G) = \frac{1}{4}(1-\delta)\left(1-k\right)\bar{t}^2 \tag{6.19}$$

Note, that at this price level, G will have no demand in the detailed market, although G's deviation price level is lower than the brand-name price level:

$$\widehat{D}^G(DM) = k\left(\frac{\bar{p}-\widehat{p}_2^G}{\theta-1}-\widehat{p}_2^G\right) = \frac{1}{2}k\frac{2\bar{p}-\bar{t}\theta}{\theta-1} > 0 \quad \Leftrightarrow \quad \bar{p} > \frac{1}{2}\bar{t}\theta \tag{6.20}$$

This can never be fulfilled, because the price cap would then no longer restrict the off-patent prices:

$$\frac{1}{2}\bar{t}\theta > \bar{t}(\theta-1)\frac{k+2\theta-1}{3k+4\theta-4} \quad \Leftrightarrow \quad \theta > 2\frac{1-k}{k+2} \tag{6.21}$$

$$\frac{1}{2}\bar{t}\theta > \bar{t}\frac{\theta}{2(k+\theta-\theta k)} \quad \Leftrightarrow \quad k < 1 \tag{6.22}$$

Both conditions are fulfilled.
G does not deviate, if $R_2^G(p_2^{G*}) \geq \widehat{R}_2^G$. This is the case, if

$$\bar{p} \geq \bar{p}(0) := \frac{\bar{t}}{k}\left[-(\theta-1)(1-k) + \sqrt{(1-k)(\theta-1)(k+\theta-1)}\right] \tag{6.23}$$

i.e. the price cap is bounded from below.

Case 2: $\bar{p} \leq \frac{1}{2}\bar{t}$

If the price cap is binding, then G optimally sets the highest possible price level that just guarantees that it can serve the whole not-detailed market: $p_2^G = \bar{p} - \epsilon$, $\epsilon \to 0$. The deviation return is then

$$\widehat{R}_2^G = (1-\delta)(1-k)\bar{p}(\bar{t}-\bar{p}) \tag{6.24}$$

G does not deviate, if $R_2^G(p_2^{G*}) \geq \widehat{R}_2^G$. This is the case, if $\bar{p} \notin [\bar{p}(3), \bar{p}(4)]$, where

$$\bar{p}(3) := (\theta-1)\,\bar{t}\,(1-k)\,\frac{k+2(\theta-1)-2\sqrt{k\,(\theta-1)\,(k+\theta-1)}}{4\theta(1-k)(\theta+k-2)+(5k^2-8k+4)} \tag{6.25}$$

$$\bar{p}(4) := (\theta-1)\,\bar{t}\,(1-k)\,\frac{k+2(\theta-1)+2\sqrt{k\,(\theta-1)\,(k+\theta-1)}}{4\theta(1-k)(\theta+k-2)+(5k^2-8k+4)} \tag{6.26}$$

This means that the price cap must be either very low or high.

Taking both the incumbent's and the entrant's incentives together, the conditions can be summarised as follows: $\bar{p} \in (p_{low}, p_{high})$ with $p_{high} = \bar{p}(2)$ and

$$p_{low} = \begin{cases} \max\{\bar{p}(0), \bar{p}(1)\} & \text{if} \quad \bar{p} > \frac{1}{2}\bar{t} \\ \max\{\bar{p}(4), \bar{p}(1)\} & \text{if} \quad \bar{p} \leq \frac{1}{2}\bar{t} \end{cases}$$

6.2 Reference Pricing

6.2.1 The Equilibrium under NRP

In a vertically separating equilibrium, characterised by a price vector (p_0, p_1, p_G), the following conditions must hold:

Condition 1: $p_G \geq 0$

Condition 2: $U_L(x, G) \geq U_L(x, 0)$

Condition 3: $U_H(x, 0) \geq U_H(x, G)$

Condition 4: $U_H(\widetilde{x}_H, 0) \geq 0$

Condition 5: $U_L(\widetilde{x}_L, G) \geq 0$

Condition 6: $\pi_0(p_0, p_1, p_G) \geq \pi_0(\widehat{p}_0, p_1, p_G)$, where $\widehat{p}_0$ solves $U_L(x, G) = U_L(x, 0)$

Condition 7: $\pi_G(p_0, p_1, p_G) \geq \pi_G(p_0, p_1, \widehat{p}_G)$, where $\widehat{p}_G$ solves $U_H(x, 0) = U_H(x, G)$

The first condition simply states that the generic price must be non-negative. The Conditions 2 and 3 ensure that the equilibrium really separates, i.e. that H-types choose the brand-name drug 0, while the L-types choose the generic substitute. Conditions 4 and 5 secure full market coverage, requiring that the indifferent patients obtain non-negative utility from purchasing and consuming either of the drugs. Finally, the Condition 6 (7) ensures that Firm 0 (Firm G) has no incentive to deviate by reducing its price and to serve the L-types (H-types).

In the following, the price equilibrium is derived in detail for the NRP case. For the two other cases, where the derivation of the equilibrium follows an identical procedure, only the constraints are presented that support the equilibrium.

The profit functions are given by (3.7) with $c_i = \alpha p_i$. Unconstrained pricing by all three firms cannot constitute an equilibrium. Unconstrained maximisation of the firms' profit functions yields the following reaction functions:

$$p_0 = \frac{1}{2\alpha}(t + \alpha p_1) \tag{6.27}$$

$$p_1 = \frac{1}{2\alpha}[t + (1-\lambda)(1-\theta)\gamma v + \alpha p_G(1-\lambda) + \alpha\lambda p_0] \tag{6.28}$$

$$p_G = \frac{1}{2\alpha}[t + \alpha p_1 - \gamma v(1-\theta)] \tag{6.29}$$

which yield the following candidate equilibrium prices:

$$p_0 = \frac{1}{\alpha}\left[t + \frac{1}{6}\gamma v\,(1-\theta)\,(1-\lambda)\right] \tag{6.30}$$

$$p_1 = \frac{1}{\alpha}\left[t + \frac{1}{3}\gamma v\,(1-\theta)\,(1-\lambda)\right] \tag{6.31}$$

$$p_G = \frac{1}{\alpha}\left[t - \frac{1}{6}\gamma v\,(1-\theta)\,(2+\lambda)\right] \tag{6.32}$$

These price vectors always violate Condition 2. In the NRP case, Condition 2 can be expressed as

$$p_G \leq p_0 - \frac{1}{\alpha}\gamma v\,(1-\theta) \tag{6.33}$$

Using (6.30) and (6.32), this condition reduces to $1 \geq 2$, which is a contradiction. In other words, (6.30)-(6.32) cannot be an equilibrium, because p_G is too high to induce even the L-type patients to buy the generic drug. Consequently, an equilibrium must be found where the generic drug is priced sufficiently low, so that not only the L-types are not induced to switch to drug 0, but firm 0 must also have no incentive to capture the L-types by lowering its price from the equilibrium level.

Using (6.27)-(6.28), the profit of firm 0 as a function of p_G is:

$$\pi_0\,(p_G) = \frac{\lambda\left[3t + (1-\lambda)\,(\alpha p_G + (1-\theta)\gamma v)\right]^2}{2\alpha t\,(4-\lambda)^2} \tag{6.34}$$

Firm 0 can drive the generic competitor out of the market and capture equal shares of the H- and L-types by setting a price

$$\widehat{p}_0 = p_G + \frac{1}{\alpha}\gamma v\,(1-\theta) \tag{6.35}$$

which yields a 'deviation' profit given by

$$\widehat{\pi}_0\,(p_G) = \frac{\left[6t - (2+\lambda)\,(\alpha p_G + (1-\theta)\gamma v)\right](\alpha p_G + (1-\theta)\gamma v)}{2\alpha t\,(4-\lambda)} \tag{6.36}$$

The optimal strategy for firm G is thus to set a price p_G that is just low enough to make such a deviation unprofitable. This price is given by the solution to

$$\pi_0\,(p_G) = \widehat{\pi}_0\,(p_G) \tag{6.37}$$

The price equilibrium can be derived by solving the three equations (6.27), (6.28) and (6.37). The solution is presented as (3.9)-(3.11) in Section 3.3.1.

It remains to specify the Conditions 1-7 for the NRP case. By construction of the equilibrium, Condition 6 is automatically satisfied. Condition 2 is also always satisfied. In the NRP case, this condition is given by

$$\theta\gamma v - \alpha p_G^{NRP} \geq \gamma v - \alpha p_0^{NRP} \tag{6.38}$$

which, using (3.9) and (3.11), reduces to

$$\Delta_0 - \Delta_G \geq 0 \tag{6.39}$$

which is true for all $\lambda \in (0,1)$. The remainder of the constraints can be expressed in the form of four different conditions on t. From (3.11), note that a non-negative generic drug price, Condition 1, is guaranteed if

$$t \geq t_1^{NRP} := \frac{(1-\theta)\,\gamma v}{3\Delta_G} \tag{6.40}$$

Furthermore, non-negative utility for the indifferent consumers of the H- and L-type, respectively, is guaranteed if

$$t \leq t_4^{NRP} := \frac{2v}{1+3\Delta_0+\Delta_1} \tag{6.41}$$

and

$$t \leq t_5^{NRP} := \frac{2\gamma v}{1+\Delta_1+3\Delta_G} \tag{6.42}$$

The necessary Condition 7 is not analytically solvable. However, to simplify, a *sufficient* condition on t can be found that satisfies Conditions 3 and 7 simultaneously. By assuming that the H-types always prefer drug 0 over drug G for the equilibrium price p_0^{NRP} and a zero-priced generic drug (i.e. $p_G = 0$), it must be true that the H-types always prefer drug 0 in equilibrium (for a non-negative generic drug price) *and* that price-undercutting by the generic firm in order to capture the H-type consumers is not an option. Using p_0^{NRP} from (3.9) and setting $p_G = 0$, this condition is given by

$$t \leq t_7^{NRP} := \frac{v\,(1-\theta)}{3\Delta_0} \tag{6.43}$$

To sum up, a price equilibrium exists in the NRP case and is given by (3.9)-(3.11), when $t \in \left[\underline{t}, \overline{t}^{NRP}\right]$, where $\underline{t} := t_1^{NRP}$ and $\overline{t}^{NRP} := \min\left\{t_4^{NRP}, t_5^{NRP}, t_7^{NRP}\right\}$. In general, existence of the equilibrium requires that the share of the L-types is relatively low combined with a sufficiently large difference in the gross valuations between the two types. To give an illustrative numerical example, assume that $v = 1$, $\lambda = 0.9$, $\theta = 0.8$ and $\gamma = 0.4$. In this

case, $\underline{t} = 0.12$ and $\overline{t}^{NRP} = t_7^{NRP} = 0.20$. Note also that the equilibrium exists for an even wider range of mismatch costs, since the upper bound $\overline{t}^{NRP}$ in this case is a sufficient, but not a necessary condition.

6.2.2 Therapeutic Reference Pricing

The price equilibrium under TRP is derived similarly to the NRP case and is given by (3.22) in Section 3.3.2. As before, the Condition 6 is automatically satisfied. Furthermore, the Conditions 1 and 2 are identical under NRP and TRP. The remainder of the Conditions – 4, 5 and 3+7 – are given by, respectively,

$$t \leq t_4^{TRP} := \frac{2\left(1-\gamma\left(1-\theta\right)\left(1-\alpha\right)\right)v}{1+3\Delta_0+\Delta_1-6\Delta_G\left(1-\alpha\right)} \tag{6.44}$$

$$t \leq t_5^{TRP} := \frac{2\left(\theta+\alpha\left(1-\theta\right)\right)\gamma v}{1+\Delta_1-3\Delta_G\left(1-2\alpha\right)} \tag{6.45}$$

$$t \leq t_7^{TRP} := \frac{\left(1-\gamma\left(1-\alpha\right)\right)\left(1-\theta\right)v}{3\left(\Delta_0-\Delta_G\left(1-\alpha\right)\right)} \tag{6.46}$$

Thus, under TRP, an equilibrium exists and is given by (3.22), when $t \in \left[\underline{t}, \overline{t}^{TRP}\right]$, where $\overline{t}^{TRP} := \min\left\{t_4^{TRP}, t_5^{TRP}, t_7^{TRP}\right\}$. It is worth noting that, due to lower equilibrium prices, the range of mismatch costs for which the equilibrium exists is generally wider under TRP. Using the same numerical example as in the NRP case, with a 10 per cent co-payment rate ($\alpha = 0.1$), the lower and upper bounds on t are given by $\underline{t} = 0.12$ and $\overline{t}^{TRP} = t_7^{TRP} = 0.34$.

6.2.3 Generic Reference Pricing

The price equilibrium under GRP is derived similarly to the NRP and TRP cases and given by (3.25)-(3.27) in Section 3.3.2. As before, Condition 6 is automatically satisfied.

Using (3.25)-(3.27), the remainder of the conditions can be derived that support the equilibrium under GRP. Once more, it can be shown that Condition 1 is satisfied, if $t \geq \underline{t}$, implying that Condition 1 is identical for all three regimes.

Condition 2 is given by

$$t \geq t_2^{GRP} := \frac{1}{3}\left(1-\alpha\right)\left(1-\theta\right)\gamma v \tag{6.47}$$

Since $t_1^{GRP} \geq (1-\theta)\gamma v$, it follows that $\underline{t} \geq t_2^{GRP}$. Thus, as long as Condition 1 is satisfied, Condition 2 is also automatically satisfied. The Conditions 4 and 5 are given by, respectively,

$$t \leq t_4^{GRP} := \frac{2\widetilde{\Delta}v + (1-\alpha)(1-\theta)\gamma v\left(3(2+\alpha)+\kappa-2\widehat{\Delta}\right)}{\widetilde{\Delta}+\overline{\Delta}+15\alpha-3\lambda\alpha(4-\alpha-\lambda)-3(1-\lambda)(2-\lambda)+3\kappa} \tag{6.48}$$

where

$$\kappa := \sqrt{1-\lambda}\left[2+\lambda-\alpha(5-2\lambda)\right]$$

and

$$t \leq t_5^{GRP} := \gamma v \frac{\widetilde{\Delta}(1+\theta)+(1-\theta)\left[(1-\alpha)[2(2+\alpha)+\varsigma]-\widehat{\Delta}(1-2\alpha)\right]}{\widetilde{\Delta}+\overline{\Delta}+3\alpha[4-\lambda(3-\alpha-\lambda)]+3\varsigma} \tag{6.49}$$

where

$$\varsigma := \sqrt{1-\lambda}\left(\lambda(2\alpha+1)-6\alpha\right)$$

Finally, the sufficient condition that simultaneously satisfies Condition 3 and Condition 7 is given by

$$t \leq t_7^{GRP} := \frac{(1-\theta)\left(\widetilde{\Delta}v+\gamma v(1-\alpha)\left(2+\alpha+\sqrt{1-\lambda}(2-\alpha(3-\lambda))-\widehat{\Delta}\right)\right)}{3\left(5\alpha+\lambda(1-\alpha)(3-\alpha-\lambda)+\sqrt{1-\lambda}(2-\alpha(3-\lambda))-2\right)} \tag{6.50}$$

Thus, under GRP, an equilibrium exists and is given by (3.25)-(3.27), when $t \in \left[\underline{t}, \overline{t}^{GRP}\right]$, where $\overline{t}^{GRP} := \min\left\{t_4^{GRP}, t_5^{GRP}, t_7^{GRP}\right\}$. Once more, due to the general price reducing effect of reference pricing, the range of mismatch costs, for which the equilibrium exists, is generally wider under the GRP system compared to the NRP case. Using the same numerical example as previously, the lower and upper bounds on t are given by $\underline{t} = 0.12$ and $\overline{t}^{GRP} = t_7^{GRP} = 0.29$.

6.2.4 Proof of Proposition 4

In equilibrium, the price difference between the two brand-name drugs is given by

$$p_1^{GRP} - p_0^{GRP} = \frac{\gamma v(1-\alpha)(1-\theta)\left[\widehat{\Delta}-4+\alpha^2-\sqrt{1-\lambda}(\lambda-\alpha^2)\right]+t\sigma}{\alpha\widetilde{\Delta}} \tag{6.51}$$

where

$$\sigma := 4\alpha+\lambda-6\alpha\lambda-3\alpha^2+2\lambda^2-\lambda^3+\alpha\lambda^2+2\alpha^2\lambda+3\sqrt{1-\lambda}\left(\lambda-\alpha^2\right)$$

By the definition of $\widehat{\Delta}$, it can easily be verified that the sum of the four terms in the square brackets in the numerator in (6.51) is positive for $\alpha \in (0,1)$ and $\lambda \in (0,1)$. The sign of the expression depends thus on the sign of σ. Once more, it is relatively straightforward to verify that $\sigma > 0$ for all $\lambda \in (0,1)$, if $\alpha < \frac{2}{3}$. Thus, $\alpha < \frac{2}{3}$ is a *sufficient* condition for $p_1^{GRP} > p_0^{GRP}$.

Regarding the equilibrium market allocations, we derive from (3.33) that

$$\widetilde{x}_L^{GRP} > \frac{1}{2} \quad \text{if} \quad t > \gamma v\,(1-\theta)\,\beta$$

where

$$\beta := (1-\alpha)\,\frac{\sqrt{1-\lambda}\,(2\alpha+\lambda)+2\alpha+\lambda\,(3-\lambda-\alpha)}{3\sqrt{1-\lambda}\,(2\alpha+\lambda)-2\alpha+\lambda+3\alpha\lambda+2\lambda^2-\lambda^3-2\alpha\lambda^2-\alpha^2\lambda}$$

It is fairly straightforward to verify that $\beta < 1$ for $\alpha \in (0,1)$ and $\lambda \in (0,1)$. This implies that $t > \gamma v\,(1-\theta)\,\beta$ (and thus $\widetilde{x}_L^{GRP} > \frac{1}{2}$) as long as Condition 1 (non-negative generic price) is satisfied.

Now consider the indifferent type-H patient. From (3.32) we can characterise $\widetilde{x}_H^{GRP}$ as a function of t in the following way:

$$\frac{\partial \widetilde{x}_H^{GRP}}{\partial t} > 0 \quad \text{for} \quad t \neq 0$$

$$\lim_{t\to 0^+}\left(\widetilde{x}_H^{GRP}\right) \to -\infty$$

and

$$\lim_{t\to-\infty}\left(\widetilde{x}_H^{GRP}\right) = \lim_{t\to\infty}\left(\widetilde{x}_H^{GRP}\right) = \vartheta$$

where

$$\vartheta := \frac{3\left(2+\alpha-\sqrt{1-\lambda}\,(2-\lambda-\alpha)\right)}{2\widetilde{\Delta}}$$

It follows that $\widetilde{x}_H^{GRP} < \frac{1}{2}$ for $t > 0$, if $\vartheta < \frac{1}{2}$ for $\alpha \in (0,1)$ and $\lambda \in (0,1)$. On the other hand, if $\vartheta > \frac{1}{2}$ for some combinations of λ and α, it must be that $\widetilde{x}_H^{GRP} > \frac{1}{2}$, if t is sufficiently high. Solving $\vartheta = \frac{1}{2}$ for α yields a function $\alpha^*(\lambda)$, such that $\vartheta < (>)\,\frac{1}{2}$, if $\alpha < (>)\,\alpha^*(\lambda)$. It is straightforward to verify that $\partial\alpha^*/\partial\lambda > 0$ and that $\alpha^* < 0$ for $\lambda < 0.54$. It follows that $\widetilde{x}_H^{GRP} < \frac{1}{2}$, if $\lambda < 0.54$, whereas, for $\lambda > 0.54$, $\widetilde{x}_H^{GRP} > \frac{1}{2}$, if λ and/or t are sufficiently high. By numerical simulations, it is also straightforward to verify that both cases, $\widetilde{x}_H^{GRP} < \frac{1}{2}$ and $\widetilde{x}_H^{GRP} > \frac{1}{2}$, can occur in equilibrium. *Q.E.D.*

6.2.5 Equilibrium Profits under GRP

Equilibrium profits under GRP are given by

$$\pi_0^{GRP} = \frac{\left(2+\alpha-\sqrt{1-\lambda}\,(2-\lambda-\alpha)\right)^2\Gamma^2\lambda}{2t\widetilde{\Delta}^2} \tag{6.52}$$

$$\pi_1^{GRP} = \frac{(3t-\Gamma)\left(\Omega-2\sqrt{1-\lambda}\,(\lambda-2\alpha+\alpha\lambda)\,\Psi\right)+t^2\left(6\sqrt{1-\lambda}\,(\lambda-2\alpha+\alpha\lambda)\,\overline{\Delta}+\Phi\right)}{2t\alpha\widetilde{\Delta}^2} \tag{6.53}$$

$$\pi_G^{GRP} = \frac{\Gamma\left(\alpha\,(2-\lambda)+\lambda\,(3-\lambda)+\sqrt{1-\lambda}\,(2\alpha+\lambda)\right)(1-\lambda)\,\Theta}{2t\widetilde{\Delta}^2} \tag{6.54}$$

where

$$\begin{aligned}
\Omega &:= 2t\omega_1+\gamma v\,(1-\alpha)\,(1-\theta)\,\omega_2\\
\omega_1 &:= 64\alpha\lambda+8\alpha^2+2\lambda^2+9\lambda^3-12\lambda^4+6\lambda^5-\lambda^6\\
&\quad -86\alpha\lambda^2-8\alpha^2\lambda+40\alpha\lambda^3+4\alpha^3\lambda-6\alpha\lambda^4+19\alpha^2\lambda^2\\
&\quad -13\alpha^2\lambda^3-4\alpha^3\lambda^2+3\alpha^2\lambda^4+2\alpha^3\lambda^3\\
\omega_2 &:= 16\alpha\lambda+8\alpha^2+26\lambda^2-41\lambda^3+26\lambda^4-8\lambda^5+\lambda^6-30\alpha\lambda^2-16\alpha^2\lambda\\
&\quad +28\alpha\lambda^3-12\alpha\lambda^4+2\alpha\lambda^5+13\alpha^2\lambda^2-5\alpha^2\lambda^3+\alpha^2\lambda^4\\
\Psi &:= \gamma v\,(1-\alpha)\,(1-\theta)\left(2\alpha+5\lambda-2\alpha\lambda-4\lambda^2+\lambda^3+\alpha\lambda^2\right)\\
&\quad +2t\left(2\alpha-7\lambda+7\lambda^2-2\lambda^3-\alpha\lambda^2+\alpha^2\lambda\right)\\
\Phi &:= 136\alpha^2-16\alpha\lambda+10\lambda^2-5\lambda^3+2\lambda^4-4\lambda^5+\lambda^6+82\alpha\lambda^2\\
&\quad -192\alpha^2\lambda-60\alpha\lambda^3+40\alpha^3\lambda+16\alpha\lambda^4-2\alpha\lambda^5+105\alpha^2\lambda^2\\
&\quad -13\alpha^2\lambda^3-24\alpha^3\lambda^2-3\alpha^2\lambda^4+4\alpha^3\lambda^3+4\alpha^4\lambda^2\\
\Theta &:= 3t\left(\alpha\lambda-3\lambda+\lambda^2+4\right)-\gamma v\,(1-\theta)\,\widehat{\Delta}-\sqrt{1-\lambda}\,(4-\lambda)\,\Gamma
\end{aligned}$$

6.2.6 Proof of Proposition 5

A direct analytical comparison of the equilibrium profits for firm 1 under the three different regimes is infeasible, since the equilibrium profit expression under GRP is extremely tedious. However, we can prove the proposition via a somewhat more subtle route, by considering, how different reimbursement systems affect the equilibrium prices and the market shares. From Proposition 3, we know that there is a clear-cut ranking of the equilibrium prices across

the different regimes, where $p_i^{NRP} > p_i^{GRP} > p_i^{TRP}$, $i = 0, 1, G$. Regarding the equilibrium market shares, we know that these are identical under NRP and TRP. Furthermore, we also know that $\widetilde{x}_j^{GRP} > \widetilde{x}_j^{TRP} = \widetilde{x}_j^{NRP}$, $j = H, L$. Thus, since $p_1^{NRP} > p_1^{GRP} > p_1^{TRP}$ and demand is at least as high under NRP than under any other reimbursement regime, it follows unambiguously that $\pi_1^{NRP} > \max\left\{\pi_1^{GRP}, \pi_1^{TRP}\right\}$. Regarding the comparison between GRP and TRP, it is not immediately obvious that firm 1 earns higher profits under GRP, since prices are higher, but market shares are lower, compared with TRP. Note, however, that the equilibrium prices are higher for *all* firms under GRP, compared with TRP. Furthermore, we know that, for given prices, $c_1^{GRP} < c_1^{TRP}$. Thus, if firm 1 unilaterally deviates from the GRP equilibrium by setting a price equal to the equilibrium price under TRP, this firm will increase its market shares, in both consumer segments, *beyond* its equilibrium market shares under TRP, and consequently earn higher profits than under TRP. Such a deviation is not profitable, so firm 1 must earn even higher profits in the GRP equilibrium, where $p_1^{GRP} > p_1^{TRP}$. *Q.E.D.*

6.2.7 Equilibrium Mismatch Costs

Inserting the expressions for the locations of indifferent patients in the different reimbursement regimes, reported throughout Section 3, into (3.36), the equilibrium mismatch costs are given by

$$C_{NRP} = C_{TRP} = \frac{\left(\delta - 6\sqrt{1-\lambda}\,(5\lambda+4)\,(1-\lambda)^3\right)t}{4\left(3\lambda - 3\lambda^2 + \lambda^3 + 8\right)^2} \tag{6.55}$$

where

$$\delta := 104 + 6\lambda - 78\lambda^2 + 53\lambda^3 - 15\lambda^4 + 15\lambda^5 - 4\lambda^6$$

and

$$C_{GRP} = \frac{(3t-\Gamma)\left[(3t-\Gamma)\left(2\sqrt{1-\lambda}\Upsilon - \Lambda\right) + 2t\left(\sqrt{1-\lambda}\mu - \eta\right)\right] + t^2\left(\tau + 6\xi\sqrt{1-\lambda}\right)}{4t\widetilde{\Delta}^2} \tag{6.56}$$

where

$$
\begin{aligned}
\xi &:= 16\alpha\lambda - 12\lambda - 4\alpha^2 + 23\lambda^2 - 17\lambda^3 + 4\lambda^4 - 18\alpha\lambda^2 + 5\alpha^2\lambda \\
&\quad +4\alpha\lambda^3 - 2\alpha^3\lambda + \alpha\lambda^4 - 3\alpha^2\lambda^2 + 2\alpha^2\lambda^3 + \alpha^3\lambda^2 \\
\tau &:= 72\lambda + 64\alpha\lambda + 104\alpha^2 - 94\lambda^2 + 74\lambda^3 - 53\lambda^4 + 25\lambda^5 - 4\lambda^6 \\
&\quad -150\alpha^2\lambda - 30\alpha\lambda^3 + 20\alpha^3\lambda + 44\alpha\lambda^4 - 10\alpha\lambda^5 + 66\alpha^2\lambda^2 \\
&\quad +7\alpha^2\lambda^3 - 12\alpha^3\lambda^2 - 6\alpha^2\lambda^4 + 2\alpha^3\lambda^3 + 2\alpha^4\lambda^2 - 40\alpha\lambda^2 \\
\Lambda &:= 12\lambda^3 - 16\alpha\lambda - 8\alpha^2 - 2\lambda^2 - 8\lambda - 7\lambda^4 + \lambda^5 + 24\alpha\lambda^2 \\
&\quad +14\alpha^2\lambda - 14\alpha\lambda^3 + 2\alpha\lambda^4 - 8\alpha^2\lambda^2 + \alpha^2\lambda^3 \\
\Upsilon &:= 8\alpha\lambda - 4\lambda + 4\alpha^2 + 5\lambda^2 - 4\lambda^3 + \lambda^4 - 10\alpha\lambda^2 - 5\alpha^2\lambda + 3\alpha\lambda^3 + 2\alpha^2\lambda^2 \\
\mu &:= 24\lambda - 40\alpha\lambda - 8\alpha^2 - 38\lambda^2 + 29\lambda^3 - 7\lambda^4 + 48\alpha\lambda^2 + 10\alpha^2\lambda \\
&\quad -13\alpha\lambda^3 + 2\alpha^3\lambda - \alpha\lambda^4 - 3\alpha^2\lambda^2 - 2\alpha^2\lambda^3 - \alpha^3\lambda^2 \\
\eta &:= 24\lambda - 8\alpha\lambda + 8\alpha^2 - 34\lambda^2 + 21\lambda^3 - 12\lambda^4 + 6\lambda^5 - \lambda^6 \\
&\quad +16\alpha\lambda^2 - 14\alpha^2\lambda - 20\alpha\lambda^3 - 2\alpha^3\lambda + 15\alpha\lambda^4 - 3\alpha\lambda^5 \\
&\quad -5\alpha^2\lambda^2 + 11\alpha^2\lambda^3 + 2\alpha^3\lambda^2 - 3\alpha^2\lambda^4 - \alpha^3\lambda^3
\end{aligned}
$$

Chapter 7

References

1. Acemoglu, Daron and Joshua Linn (2003), *Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry*, NBER Working Paper 10038

2. Arfwedson, Jacob (2004), *Parallel Trade in Pharmaceuticals*, IPI Policy Report 182

3. Aronsson, Thomas, Mats A. Bergman, and Niklas Rudholm (2001) *The Impact of Generic Drug Competition on Brand Name Market Shares – Evidence from Micro Data*, Review of Industrial Organization 19: 425-435

4. Arrow, Kenneth J. (1963), *Uncertainty and the Welfare Economics of Medical Care*, The American Economic Review 53(5): 941-973

5. Atkinson, Timothy J. (2001), *The Global Parallel Trade Outlook 2001-2006: A Country-by-Country Analysis*, Reuters Business Insight

6. Avorn, Jerry, Milton Chen, and Robert Hartley (1982), *Scientific vs. Commercial Sources of Influence on the Prescribing Behaviors of Physicians*, American Journal of Medicine 73(1): 4-8

7. Bale, Harvey E. Jr. (1998), *The Conflicts between Parallel Trade and Product Access and Innovation: The Case of Pharmaceuticals*, Journal of International Economic Law 1: 637-653

8. Ballance, Robert, Janos Pogany, and Helmut Forstner (1992), *World's Pharmaceutical Industries: An International Perspective on Innovation, Competition and Policy*,

Edward Elgar Publishing, Cheltenham, UK

9. Berndt, Ernst R. (2002), *Pharmaceuticals in U.S. Health Care: Determinants of Quantity and Price*, Journal of Economic Perspectives 16(4): 45-66

10. Bordoy, Catalina and Izabela Jelovac (2005), *Pricing and Welfare Implications of Parallel Imports in the Pharmaceutical Industry*, International Journal of Health Care Finance and Economics 5: 5-21

11. Brander, James A. and Barbara J. Spencer (1985), *Export Subsidies and International Market Share Rivalry*, Journal of International Economics 18: 83-100

12. Brekke, Kurt R. and Michael Kuhn (2006), *Direct-to-Consumer Advertising in Pharmaceutical Markets*, Journal of Health Economics 25: 102-130

13. Brekke, Kurt R., Robert Nuscheler, and Odd Rune Straume (2006), *Quality and Location Choices under Price Regulation*, Journal of Economics and Management Strategy 15(1): 207-228

14. Breyer, Friedrich, Peter Zweifel, and Mathias Kifmann (2003), *Gesundheitsökonomie*, Berlin, Heidelberg, New York: Springer-Verlag

15. Cabrales, Antonio (2003), *Pharmaceutical Generics, Vertical Product Differentiation, and Public Policy*, Working Paper UPF 662

16. Caves, Richard E., Michael D. Whinston, and Mark A. Hurwitz (1992), *Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry*, Brookings Papers on Economic Activity: 1-48

17. Chard, J.S. and C.J. Mellor (1989), *Intellectual Property Rights and Parallel Imports*, The World Economy 12: 69-83

18. Chren, Mary-Margaret and C. Seth Landefeld (1994), *Physicians' Behavior and their Interactions with Drug Companies*, Journal of the American Medical Association 271(9): 684-689

19. Danzon, Patricia M. (1997a), *Pharmaceutical Price Regulation*, Washington, DC: The American Enterprise Institute Press

20. Danzon, Patricia M. (1997b), *Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU*, International Journal of the Economics of Business 4(3): 301-321

21. Danzon, Patricia M. (1998), *The Economics of Parallel Trade*, Pharmacoeconomics 13(3): 293-304

22. Danzon, Patricia M. (2001), *Reference Pricing: Theory and Evidence*, In: Lopez-Casasnovas, G. and B. Jönsson (eds.): Reference Pricing and Pharmaceutical Policy: Perspectives on Economics and Innovation, Springer, Barcelona

23. Danzon, Patricia M. and LiWei Chao (2000), *Does Regulation Drive Out Competition in Pharmaceutical Markets?*, Journal of Law and Economics 43: 311-357

24. Danzon, Patricia M. and Jonathan D. Ketcham (2004), *Reference Pricing of Pharmaceuticals for Medicare: Evidence from Germany, the Netherlands and New Zealand*, In: Cutler, D.M. and A.M. Garber (eds.): Frontiers in Health Policy Research, vol. 7, National Bureau of Economic Research and MIT Press

25. Danzon, Patricia M. and Hong Liu (1998), *Reference Pricing and Physician Drug Budgets: The German Experience in Controlling Pharmaceutical Expenditures*, Working Paper, Wharton School

26. Danzon, Patricia M. and Adrian Towse (2003), *Differential Pricing for Pharmaceuticals: Reconciling Access, R&D and Patents*, Working Paper 03-7, AEI-Brookings Joint Center for Regulatory Studies

27. Danzon, Patricia M., Y. Richard Wang, and Liang Wang (2003), *The Impact of Price Regulation on the Launch Delay of new Drugs – Evidence from Twenty-Five Major Markets in the 1990s*, NBER Working Paper 9874

28. Darba, Josep and Joan Rovira (1998), *Parallel Imports of Pharmaceuticals in the European Union*, Pharmacoeconomics 14, Suppl. 1: 129-136

29. DiMasi, Joseph A., Ronald W. Hansen, Henry G. Grabowski, and Louis Lasagna (1991), *Cost of Innovation in the Pharmaceutical Industry*, Journal of Health Economics 10: 107-142

30. DiMasi, Joseph A., Ronald W. Hansen, Henry G. Grabowski, and Louis Lasagna (2003), *The Price of Innovation: New Estimates of Drug Development Costs*, Journal of Health Economics 22: 151-185

31. Dixit, Avinash and Victor Norman (1978), *Advertising and Welfare*, Bell Journal of Economics 9: 1-17

32. Economides, Nicholas (1989), *Quality Variations and Maximal Variety Differentiation*, Regional Science and Urban Economics 19: 21-29

33. Frank, Richard G. and David S. Salkever (1992), *Pricing, Patent Loss and the Market for Pharmaceuticals*, Southern Economic Journal 59(2): 165-179

34. Frank, Richard G. and David S. Salkever (1997), *Generic Entry and the Market for Pharmaceuticals*, Journal of Economics & Management Strategy 6: 75-90

35. Fridman, D., A. Jaffe and S. Steinhardt (1987), *Physicians' Attitudes toward and Knowledge about Generic Drug Substitution*, New York State Journal of Medicine 87: 539-542

36. Fudenberg, Drew and Jean Tirole (1984), *The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look*, American Economic Review Papers and Proceedings 74: 361-366

37. Gabszewicz, Jean J. and Jacques F. Thisse (1979), *Price Competition, Quality and Income Disparities*, Journal of Economic Theory 20(3): 340-359

38. Gabszewicz, Jean J. and Jacques F. Thisse (1980), *Entry (and Exit) in a Differentiated Industry*, Economic Journal 96: 160-172

39. Gambardella, Alfonso (1995), *Science and Innovation: The US Pharmaceutical Industry During the 1980s*, Cambridge University Press

40. Grabowski, Henry G. and John M. Vernon (1990), *A New Look at the Returns and Risks to Pharmaceutical R&D*, Management Science 36: 804-821

41. Grabowski, Henry G. and John M. Vernon (1992), *Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act*, Journal of Law and Economics 35: 331-350

42. Grabowski, Henry G. and John M. Vernon (2000), *The Determinants of Pharmaceutical Research and Development Expenditures*, Journal of Evolutionary Economics 10: 201-215

43. Hassett, Kevin A. (2004), *Price Controls and the Evolution of Pharmaceutical Markets*, American Enterprise Institute for Public Policy Research

44. Hellerstein, Judith K. (1998), *The Importance of the Physician in the Generic versus Trade-Name Prescription Decision*, RAND Journal of Economics 29(1): 108-136

45. Hudson, John (2000), *Generic Take-Up in the Pharmaceutical Market Following Patent Expiry: A Multi-Country Study*, International Review of Law and Economics (20): 205-221

46. Hotelling, Harold (1929), *Stability in Competition*, Economic Journal 39: 41-57

47. Hurwitz, Mark A. and Richard E. Caves (1988), *Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals*, Journal of Law and Economics 31: 299-320

48. Huskamp, Haiden A., Meredith B. Rosenthal, Richard G. Frank, and Joseph P. Newhouse (2000), *The Medicare Prescription Drug Program: How will the Game be Played?*, Health Affairs 19(2): 18-23

49. Ishigaki, Hiroaki (2000), *Informative Advertising and Entry Deterrence: A Bertrand Model*, Economic Letters 67: 337-343

50. Jacobzone, Stephane (2000), *Pharmaceutical Policies in OECD Countries: Reconciling Social and Industrial Goals*, OECD Labour Market and Social Policy Occasional Papers No. 40

51. Kanavos, Panos (1998), *The Politics of Health and the Industrial Economics of Pharmaceuticals: Is there a Single European Dimension?*, Political Studies Association Conference Proceedings 1998

52. Kanavos, Panos (2002), *Pharmaceutical Regulation in Europe*, IRPP Conference Paper

53. Kanavos, Panos and Uwe Reinhardt (2003), *Reference Pricing for Drugs: Is it Compatible with U.S. Health Care?*, Health Affairs 22(3): 16-30

54. Königbauer, Ingrid (2005), *Advertising and Generic Market Entry*, HEB Working Paper 03/2005, University of Bergen

55. Krugman, Paul R. (1984), *Import Protection as Export Promotion: International Competition in the Presence of Oligopoly and Economies of Scale*, In: Kierzskowski, H. (eds.): Monopolistic Competition and International Trade, Oxford University Press, Oxford

56. Leffler, Keith B. (1980), *Persuasion or Information? The Economics of Prescription Drug Advertising*, Journal of Law and Economics 24: 45-74

57. Lichtenberg, Frank R. (2001), *Are the Benefits of New Drugs Worth Their Cost?*, Health Affairs 20: 241-251

58. Lopez-Casasnovas, Guillem and Jaume Puig-Junoy (2000), *Review of the Literature on Reference Pricing*, Health Policy 54: 87-123

59. Lundin, Douglas (2000), *Moral Hazard in Physician Prescription Behavior*, Journal of Health Economics 19: 639-662

60. Ma, Ching-to Albert and James F. Burgess Jr. (1993), *Quality Competition, Welfare, and Regulation*, Journal of Economics 58(2): 153-173

61. Malueg, David A. and Marius Schwartz (1994), *Parallel Imports, Demand Dispersion, and International Price Discrimination*, Journal of International Economics 37: 167-195

62. Maskus, Keith E. (2000), *Parallel Imports*, The World Economy: Global Trade Policy 2000 23(9): 1269-1284

63. Merino-Castelló, Anna (2003), *Impact of the Reference Price System on the Pharmaceutical Market: A Theoretical Approach*, Working Paper Ref. 524, Universitat Pompeu Fabra

64. Mushkin, Selma J. (1958), *Toward a Definition of Health Economics*, Public Health Reports 73(9): 785-793

65. OECD Health Data (2003), *OECD Data Show Health Expenditures at an All-time High*, Download: www.oecd.org/document/39/0,2340,en_2825_495642_2789735_1_1_1_1,00.html (11. Mai 2004)

66. Pavcnik, Nina (2002), *Do Pharmaceutical Prices respond to Potential Patient Out-of-Pocket Expenses?*, RAND Journal of Economics 33(3): 469-487

67. Pecorino, Paul (2002), *Should the U.S. Allow Prescription Drug Reimports from Canada?*, Journal of Health Economics 21: 699-708

68. Rizzo, John A. (1999), *Advertising and Competition in the Ethical Pharmaceutical Industry: The Case of Antihypertensive Drugs*, Journal of Law and Economics 42: 89-116

69. Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen (2005), *Gutachten 2005: Koordination und Qualität im Gesundheitswesen*, Download: www.svr-gesundheit.de (19. December 2005)

70. Scherer, F. Michael (1993), *Pricing, Profits and Technological Progress in the Pharmaceutical Industry*, Journal of Economic Perspectives 7: 97-115

71. Scherer, F. Michael (2000), *The Pharmaceutical Industry*, In: Culyer, A.J. and J.P. Newhouse (eds.): Handbook of Health Economics, Volume IB, Elsevier, Amsterdam

72. Schmalensee, Richard (1983), *Advertising and Entry Deterrence: An Exploratory Model*, Journal of Political Economy 91(4): 636-653

73. Scott Morton, Fiona M. (2000), *Barriers to Entry, Brand Advertising, and Generic Entry in the U.S. Pharmaceutical Industry*, International Journal of Industrial Organization 18(7): 1085-1104

74. Shaked, Avner and John Sutton (1982), *Relaxing Price Competition Through Product Differentiation*, Review of Economic Studies 99: 3-13

75. Shaked, Avner and John Sutton (1983), *Natural Oligopolies*, Econometrica 51(5): 1469-1483

76. Taggart, James (1993), *The World Pharmaceutical Industry*, Thomson Learning, London: Routledge

77. Telser, Lester G., William Best, John W. Egan, and Harlow N. Higinbotham (1975), *The Theory of Supply with Applications to the Ethical Pharmaceutical Industry*, Journal of Law and Economics 18(2): 449-478

78. Vernon, John M. (1971), *Concentration, Promotion, and Market Share Stability in the Pharmaceutical Industry*, Journal of Industrial Economics 19(3): 246-266

79. Zweifel, Peter and Luca Crivelli (1996), *Price Regulation of Drugs: Lessons from Germany*, Journal of Regulatory Economics 10: 257-273

Lebenslauf

Persönliche Angaben

Name:	Ingrid Königbauer
Geburtsdatum:	25. Januar 1977
Geburtsort:	Pfarrkirchen
Staatsangehörigkeit:	Deutsch

Schulischer Werdegang

1983 - 1987	Grundschule Pfarrkirchen
1987 - 1996	Gymnasium Pfarrkirchen
	Mai 1996: Abitur

Akademischer Werdegang

1997 - 2002	Studium der Volkswirtschaftslehre
	Ludwig–Maximilians-Universität München
	Juni 2002: Diplom-Volkswirtin
1999 - 2000	Undergraduate Studies of Economics
	University of Warwick, UK
seit 09 / 2002	Wissenschaftliche Mitarbeiterin
	Seminar für Versicherungswissenschaft
	Ludwig–Maximilians-Universität München
2004 - 2005	Forschungsaufenthalt
	University of Bergen, Norwegen

6. März 2006